173

DEC 2- 1936

Vol. XIX, No. 6 *PSYCHOLOGICAL REVIEW PUBLICATIONS* December, 1936

Journal of Experimental Psychology

EDITED BY

SAMUEL W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA

WALTER S. HUNTER, BROWN UNIVERSITY (*Psychol. Index*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY (*Psychol. Rev.*)
JOHN A. McGEOCH, WESLEYAN UNIVERSITY (*Psychol. Bull.*)
JOHN F. DASHIELL, UNIV. OF NORTH CAROLINA (*Psychol. Monog.*)

CONTENTS

Directional Variations of Apparent Movement: C. F. WILLEY 635
Studies in Retroactive Inhibition: VII. Retroactive Inhibition as a Function of the Length and Frequency of Presentation of the Interpolated Lists: J. A. McGEOCH 674
The Effect of Practice on Positive Time-Order Errors: H. WOODROW AND L. H. STOTT 694
The Effect of Interpolated Time Intervals upon the Contrast Effects: M. G. PRESTON 706
An Experimental Isolation of Some Factors Determining Response to Rhythmic Cutaneous Stimulation. I. Frequency, Pressure and Time: R. M. BELLOWS 716
Reliability and Validity of Eye-Movement Measures of Reading: M. A. TINKER 732
The Illusion of the Oblique Intercept: P. SCHILDER AND D. WECHSLER 747
Maze Learning with a Differential Proprioceptive Cue: L. F. CARTER 758
The Influence of Degree of Wholeness on Whole-Part Learning: M. V. SEAGOE 763
A Study of the Relative Amounts of Forward and Backward Association of Verbal Material: T. G. HERMANS 769
Distribution of Effort in a Perceptual-Motor Task. Series II: D. RANSON AND E. B. SKAGGS 776
The New Laboratory of Psychology at the University of Rochester: L. CARMICHAEL 783

PUBLISHED BI-MONTHLY
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter, July 1, 1920, at the post-office at Lancaster, Pa.

PUBLICATIONS

OF THE

AMERICAN PSYCHOLOGICAL ASSOCIATION

EDITED BY

S. W. FERNBERGER, University of Pennsylvania (*J. Exper. Psych.*)
WALTER S. HUNTER, Brown University (*Abstracts*)
HENRY T. MOORE, Skidmore College (*J. Abn. and Soc. Psychol.*)
HERBERT S. LANGFELD, Princeton University (*Review*)
JOHN A. McGEOCH, Wesleyan University (*Bulletin*)
JOHN F. DASHIELL, Univ. of North Carolina (*Monographs*)

HERBERT S. LANGFELD, Business Editor

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly (10 numbers), the annual volume comprising about 720 pages. Special issues of the Bulletin consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bi-monthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 900 pages.

PSYCHOLOGICAL INDEX

(discontinued after No. 42, June, 1936)

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The Index has been published annually from 1895 to 1936, the volumes covering the literature of 1894 through 1935. Its bibliographical services will be carried on by the Psychological Abstracts.

PSYCHOLOGICAL ABSTRACTS

appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.

PSYCHOLOGICAL MONOGRAPHS

consists of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The Monographs appear at irregular intervals and are gathered into volumes of about 500 pages.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

appears quarterly, June, September, December, March, the four numbers comprising a volume of 448 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.

ANNUAL SUBSCRIPTION RATES

Review: \$5.50 (Foreign, \$5.75). **Journal:** \$7.00 (Foreign, \$7.25).
Monographs: \$6.00 per volume (Foreign, \$6.30).
Bulletin: \$6.00 (Foreign, \$6.25). **Abstracts:** \$6.00 (Foreign, \$6.25).
Abnormal and Social: \$5.00 (Foreign, \$5.25). Single copies \$1.50.
Current numbers: **Journal,** \$1.25; **Review,** \$1.00; **Abstracts,** 75c; **Bulletin,** 60c.

COMBINATION RATES

Review and Bulletin: \$10.00 (Foreign, \$10.50).
Review and J. Exp.: \$11.00 (Foreign, \$11.50).
Bulletin and J. Exp.: \$12.00 (Foreign, \$12.50).
Review, Bulletin, and J. Exp.: \$16.00 (Foreign, \$16.75).

Subscriptions, orders, and business communications should be sent to the

PSYCHOLOGICAL REVIEW COMPANY

PRINCETON, NEW JERSEY

Journal of
Experimental Psychology

VOL. XIX, No. 6 DECEMBER, 1936

DIRECTIONAL VARIATIONS OF APPARENT MOVEMENT [1]

BY CLARENCE F. WILLEY

Princeton University

The phenomena of apparent movement have long been recognized to be of peculiar significance to a central problem of perception—the subjective unity of the objectively discrete. A large number of investigations have been directed toward the discovery of basic principles to which the facts of apparent movement can be satisfactorily related. The difficulty of accomplishing such a systematization is attested by the numerous controversies which have centered about the various 'psychological' and 'physiological' theories which have been advanced to explain apparent movement. That aspect of apparent movement with which the present study is concerned, namely, the variations in direction which occur when there are two possible paths of movement, is particularly insusceptible of explanation in terms of the objective conditions of stimulation. In spite of this fact, directional variations are in several respects better suited to experimental investigation than other aspects of apparent movement. Reports of direction of movement can be made with greater ease and confidence than reports on the quality of movement, since it is not necessary to depend to the same degree on criteria which are difficult to define and probably inconstant.

The principal determinants of directional variations of apparent movement were indicated by Wertheimer (7, 218).

[1] From the Psychological Laboratory of Princeton University.

If the two lines which together form an acute angle are exposed in succession, with the times of exposure and the interval between these properly chosen, there results an impression of one line moving across the angle. If one of these lines is doubled in length (so that two supplementary angles are formed) and the shorter line is exposed first, movement will take place across the smaller of the two angles. This tendency of movement to follow the shortest of two possible paths will be referred to as 'the factor of shortest distance.' If the size of the smaller angle is gradually increased in successive presentations, the movement continues to take place in the same direction; it crosses the original sector and whatever has been added by the successive enlargements of it. The line appears to move in the same direction even though it must cross an angle of much more than 90 degrees. The tendency to the continuance of a given kind of movement will be called 'the factor of perseveration.'

DeSilva (1, 486) gave considerable attention to variations of direction in his experimental study of apparent movement. He presented two diameters in succession, changing in small steps the size of the angles separating the lines. His subjects showed a decided preference for movement in a vertical direction over movement in a horizontal direction. Only when the stimulus lines were nearly vertical would movement pass through the vertical plane of the line of regard. This preference of a majority of subjects for vertical movement has been confirmed recently by Kelly and Miles (**3**, 477). Some of DeSilva's subjects exhibited a bias for clockwise movement as well as for vertical movement.

Waals and Roelofs (**6**, 277) projected the two arms of a cross on a screen in such a way that they bisected each other. They reported that a reversal of the direction of apparent rotation followed a reversal of the order in which the lights were presented. Movement continued to cross the same sector but from the opposite direction.

Certain factors influencing the directional changes of apparent movement must frequently be in conflict. If one presents a stimulus arrangement consisting of two oblique

diameters, at right angles to each other, most observers will perceive vertical movement. If the path over which vertical movement occurs is now lengthened by changing the angles between the lines, a conflict arises between the tendency to see the movement as vertical and the tendency to see it across the shortest distance. Certain characteristic atypical perceptions are the phenomenal representation of such conflicts.

Although the existence of individual differences in the perception of directional variations of apparent movement has been mentioned by many investigators, few quantitative comparisons have been reported. Wertheimer (**7**, 219) stated that individual differences in set are regular and measurable. Waals and Roelofs (**6**, 278) found that their subjects differed in respect to the degree to which the direction of apparent rotation was compulsory. Neuhaus (**4**, 415) found wide individual differences in susceptibility to the factor of perseveration. Von Schiller (**5**, 192) reported that subjects were influenced unequally by differences of form and color. Some observers perceived movement more easily between stimulus members which were alike in form, while others were more likely to perceive movement between objects of the same color.

The range of individual differences in judgments of the direction of apparent movement justifies a more intensive consideration of these than they have so far received. The purpose of the present experiment was to investigate further the interaction of the factors which give rise to conflicts in the perception of direction of apparent movement, with particular regard to the nature and extent of individual differences in perception.

Method.—The subjects were presented with a large number of variations of an arrangement of four stimulus lights. The experiment was conducted in a dark room in which the stimulus lights and a fixation point were the only objects visible to the observer. The stimulus lights were 3.8 volt flashlight lamps enclosed in short sections of brass tubing. The end of the tube toward the observer was covered with one thickness of white adhesive tape; this gave a uniformly illuminated surface 1.5 cm. in diameter. The lights were mounted on two cross-arms, two lights on each arm, equidistant from a central axis about which the arms could be rotated. Five positions of the arms in relation to each other were used, as shown in Fig. 1. Any of these five arrangements could be rotated a total of 90 degrees to the left in eight steps of ap-

proximately 11.25 degrees. Several of the stimulus arrangements resulting from a rotation of the 'A3X' arrangement of Fig. 1 are shown in Fig. 2.

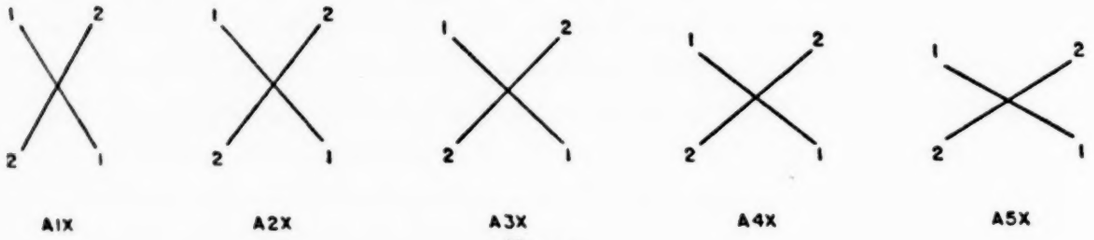


FIG. 1.

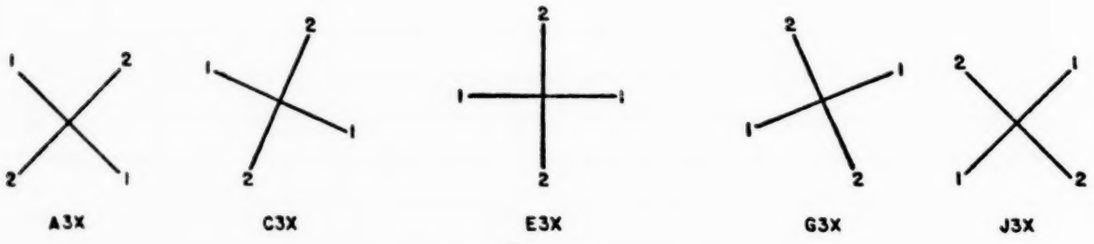


FIG. 2.

For convenient reference, the stimulus arrangements are designated by expressions consisting of a capital letter and a numeral followed by 'X' or 'Y.' The first capital letter indicates the degree of rotation to the left. The following table shows what capital letter corresponds to a given number of degrees of rotation.

A	No rotation	E	45 degrees
B	11 degrees	F	56 "
C	23 "	G	68 "
D	34 "	H	79 "
		J	90 degrees

The Arabic numeral following the first capital letter refers to one of the five stimulus arrangements of Fig. 1. The distances between the lights in each of the five positions are given in the following table. The distances are measured from the center of each light.

Arrangement of Lights	Horizontal Distance	Vertical Distance
1	4.2 cm.	7.4 cm.
2	5.5 "	6.7 "
3	6.1 "	6.1 "
4	6.7 "	5.5 "
5	7.4 "	4.2 "

In Part I the distances between the lights were slightly greater than those given above. They were subsequently reduced to improve the quality of the movement.

The capital letter 'X' or 'Y' following the Arabic numeral indicates the order in which the lights flash on. Both in Fig. 1 and in Fig. 2 only the 'X' order is shown. The lights which occupy the positions designated by '1' appear first. Those lights which are presented first with the 'X' order are presented last with the 'Y' order. The 'A3X' pattern is thus identical with the 'J3Y' pattern.

A Lewin-Schumann timing apparatus controlled the length of the exposure times and the intervals between them. An extra set of switches permitted a change in the exposure times during a series of observations without interruption of the experiment. Exposure times of 150 and 500 sigma were used in Part V of the experiment; in all the other parts the exposure time was 150 sigma. The second pair of lights appeared immediately after the first lights disappeared, so that the interval between the exposures was always zero. Between presentation of the second pair of lights and the first pair of the succeeding series there was an interval of nine seconds. During this time the experimenter could reverse the order in which the lights came on, change the pattern of the lights, or rotate the lights, without the knowledge of the observer.

The observer sat at a table so placed that his head was 450 cm. from the stimulus lights. He was instructed to fixate a small point of light at the intersection of the two cross-arms. It has been found that directing fixation or attention on a point within a path of movement makes it more likely that movement will occur within that path (**1**, 489; **6**, 278); by holding fixation on a point which was equidistant from all the lights, the influence of this factor was minimized.

The observers were graduate students in psychology or members of the staff of the department of psychology.

PART I

Procedure.—In preliminary tests the 'A3' stimulus situation yielded vertical perceptions of movement for the four subjects used. This result indicated that this pattern, like those used by DeSilva and by Kelly and Miles, favored vertical movement. Group tests of undergraduates confirmed this, as a majority of the observers saw movement as vertical when the distances between the lights were equal. The procedure of Part I was designed to determine whether or not the vertical perceptions would continue when the factor of shortest distance was so introduced as to favor horizontal perceptions. Fifteen of each of the following stimulus arrangements were given during each period of observation in this order: A3X, A2X and A1X.

Results.—Two of the subjects who observed in this part of the experiment changed from vertical to horizontal perceptions at some point during each change of the stimulus arrangements from the 'A3X' pattern through the 'A2X' to the 'A1X' pattern. They exhibited very different degrees of susceptibility, however, to the factor of shortest distance. One of these observers, Hu, generally changed from vertical to horizontal perceptions in the early part of the 'A2X' series, while the other, Ta, reported a corresponding proportion of horizontal perceptions only for the 'A1X' series. Another subject, Ka, reported no horizontal perceptions for any of the stimulus arrangements. This last result is particularly significant in view of the fact that in a previous experiment on apparent movement in which Ka had been an observer, horizontal movement between two lights was the

only possible perception. Such a circumstance creates serious difficulties for an explanation of movement perception in terms of previous experience. Certain other results obtained in Part I, however, show that the influence of experience can by no means be unconditionally excluded. Four subjects who had not observed in the preliminary tests saw movement consistently as horizontal. Since three of these subjects did not see movement of any sort until the 'A2' and 'A1' situations had been presented to them, this result may reasonably be attributed to the influence of the factor of shortest distance. This factor, as introduced in the procedure of Part I, favored the appearance of horizontal movement.

It is possible to relate the results of Part I by assuming that a vertical (or horizontal) bias exists in different degrees of intensity in different observers. If a vertical (or horizontal) bias is absent or weak, the observer may build up a tendency to whatever type of movement the conditions of stimulation favor. The four observers whose perceptions were uniformly horizontal in Part I may have had a horizontal bias. Results obtained in later parts of the experiment, however, render it more probable that they would have exhibited a weak vertical bias or none at all if the conditions of Part I had not favored horizontal movement. If a vertical bias is present and of moderate intensity, it will control the direction of movement when the equidistant stimulus arrangement, 'A3,' is presented, but it will not be able to control the 'A1' and 'A2' stimulus arrangements, where the factor of shortest distance is effective. A vertical bias of this magnitude characterizes observers Ta and Hu. Since the latter shifted to horizontal perceptions earlier in the series than Ta, his vertical bias was the weaker of the two. If a vertical bias is of still greater intensity, it will control all the perceptions of the observer, and the factor of shortest distance as introduced in Part I will have no effect. Observer Ka has such a bias; he was uninfluenced by the factor of shortest distance in Part I in spite of the fact that he had observed horizontal movement between two lights over an extended period in a previous experiment. Further evidence to support the hy-

pothesis of different intensities of perceptual bias is found in later parts of the experiment.

PART II

Procedure.—The purpose of the procedure used in Part II was to determine how the perception of movement would change with the following variations in the conditions of stimulation. Eight presentations of each of the following stimulus arrangements were given: E3Y, E3X, E3Y, E3X, C3Y, C3X, C3Y, C3X, A3Y, A3X. As the '3' stimulus arrangement (see Fig. 1) was used for all presentations in this part of the experiment, the distances between the lights were always the same. The 'E3X' arrangement is the 'A3X' rotated 45 degrees to the left, and the 'C3X' is the 'A3X' rotated 23 degrees to the left.

Results.—1. The 'E3' stimulus arrangement is less successful in eliciting rotary movement (clockwise or counter-clockwise) than the 'C3' and 'A3' arrangements. Three subjects perceived no rotary movement from the 'E3' stimulus arrangement. Kelly and Miles (**3**, 477) state that a similar stimulus pattern used by them "ordinarily gives no impression of a rotating pattern, but rather the appearance of movement toward the center of the field and then at right angles toward the periphery." Only one of the subjects who did not perceive rotary movement when the 'E3' arrangement was presented obtained non-rotary movement of this sort; the other subjects perceived virtually no movement of any kind.

Waals and Roelofs (**6**, 278), from their experiment in which rotary movement was obtained by successive presentations of the vertical and horizontal arms of a cross, concluded that the perception of rotary movement is dependent upon a proper selection of exposure times and intervals. It is clear from the experiment of Kelly and Miles, and from the present experiment, that some additional factor is involved, since the exposure times and intervals that give good movement when the square defined by the stimulus lights rests on its side no longer do so when it is rotated 45 degrees.[2]

[2] That two squares of the same size in different positions have by no means equivalent effects in perceptual experiments was shown by Hartmann (**2**, 361). He determined the maximal interval between successive brief exposures of a figure which would give complete fusion of the two images. This interval was larger for a square lying on its side than it was for a square of the same size rotated 45 degrees.

The complexity of the determinants of directional variations of apparent movement does not permit definition of these in terms of exposure times and intervals alone.

2. The three 'vertical' subjects of Part I reported widely varying perceptions for the 'E3' stimulus arrangement in Part II. Observer Ta almost without exception saw movement as counter-clockwise. Hu showed some preference for movement in an oblique direction from upper right to lower left, and observer Ka obtained no rotary movement of any sort. These three subjects saw only vertical movement when the 'A3' stimulus arrangement was presented, as in Part I.

3. It was assumed in the hypothesis advanced in the discussion of Part I that the four observers who saw only horizontal movement owed the establishment of that perceptual tendency to the fact that the conditions of stimulation favored that type of perception. Since only the 'X' order of presentation of the lights was used in Part I, the horizontal perceptions of these subjects were necessarily clockwise. The introduction in Part II of the 'Y' order of presentation for the 'C3' and 'A3' arrangements created a dilemma for these subjects. It became no longer possible to see movement both as horizontal and clockwise. The resulting conflict between the tendency to see the movement as clockwise and the tendency to see it as horizontal is apparent in the reports of these subjects. The tendency to see movement as clockwise dominated from the first, however, and at the end of five periods of observation none of the four 'clockwise' subjects perceived horizontal movement with presentation of the 'A3Y' or 'C3Y' arrangements. For these arrangements their perceptions were always clockwise-vertical; for the 'A3X' and 'C3X' arrangements, clockwise-horizontal.

A conflict of this sort between perceptual tendencies gives rise to a variety of atypical perceptions. The observer was sometimes confused and unable to make a report on the direction of movement although he was certain that some kind of movement had occurred. Sometimes there was movement from one member of the first pair of lights to both mem-

bers of the second pair. The other member of the first pair of lights was not seen at all, or was seen as motionless and independent. Sometimes there was movement over three of the four possible paths, as when clockwise movement took place over two paths and counter-clockwise over a third, with no movement over the fourth. Very rarely there was movement over all four possible paths—simultaneous clockwise and counter-clockwise movement. Sometimes the lights flashed on and off without movement, as in the successive stage of movement between two lights.

These irregular perceptions show the importance of the determinants of individual variations in the perception of direction of apparent movement. The same conditions which consistently give most observers an unequivocal type of movement gives other observers a variety of atypical perceptions such as those described above.

Part III

Procedure.—In Part III only the 'A' patterns of stimuli are used. This represents a return to the conditions of stimulation of Part I with two important modifications: (1) both the 'X' and the 'Y' orders of presentation of the lights were used, and (2) additional stimulus arrangements, 'A4' and 'A5,' were introduced. Four presentations of each of the following stimulus patterns were given in this order: A3Y, A3X, A1Y, A1X, A2Y, A2X, A3Y, A3X, A4Y, A4X, A5Y, A5X, A5Y, A5X, A4Y, A4X, A3Y, A3X, A2Y, A2X, A1Y, A1X, A3Y, A3X.

Results.—The perceptions of five subjects were relatively unaffected by the changes of the stimulus patterns in Part III, so no graphical representation of their results is given. Two of these, Wv and Ka, were always 'vertical'; two others, Ta and Hu, had only a small number of horizontal perceptions. Another observer, Cu, perceived only clockwise movement. The perceptions of these subjects show how a strong perceptual bias may overcome the factor of shortest distance.

The results for those subjects whose perceptions showed the greatest variation are shown in Figs. 3 to 5. In Part I and Part II, four subjects, Wd, Ba, Br and Cu, exhibited a striking similarity in their perceptions. These observers were consistently 'clockwise-horizontal' in Part I, and in

Part II they showed a decided preference for clockwise movement. In Part III, however, these four subjects are quite different from one another. One of these, Cu, as already noted, continued to manifest a clockwise bias which was sufficiently strong to prevent the appearance of any counter-clockwise perceptions. Ba, on the other hand, shifted from

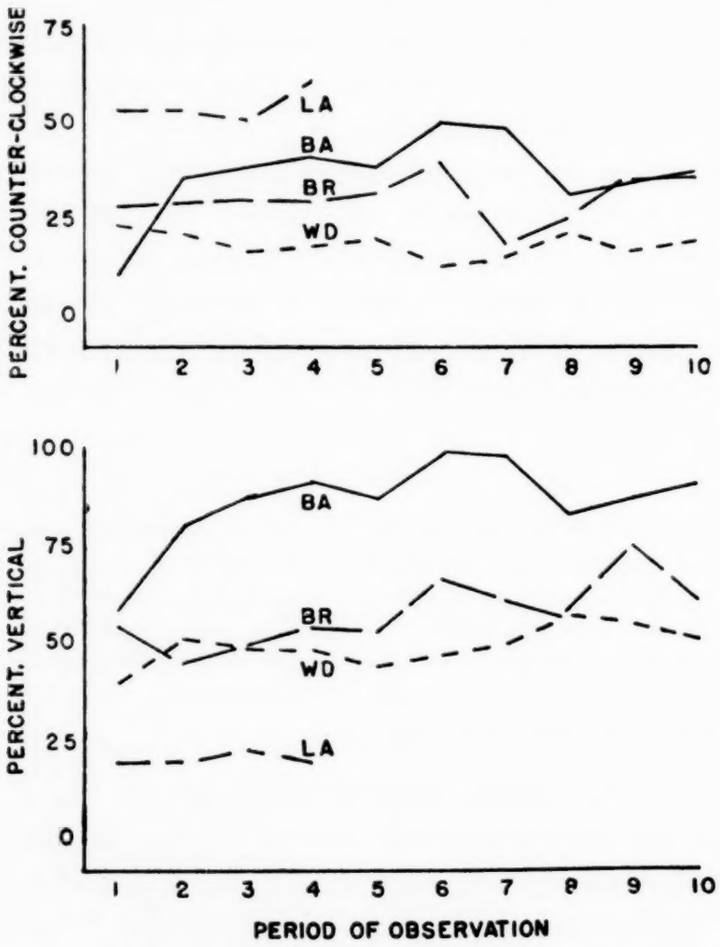


FIG. 3. Percentage of vertical and counter-clockwise perceptions reported during each period of observation in Part III.

clockwise to vertical perceptions for all stimulus arrangements except 'A1' and 'A2.' This change of dominant perceptual tendency is shown in Fig. 5 by the rise in the curve representing the perceptions of Ba. After the second day of observation in Part III his perceptions were always vertical when the 'A3' stimulus arrangement was presented. Fig. 5 brings out the contrast between the results of Ba and those of Wd

and Cu for the 'A3' pattern. The perceptions of Br also shifted to the vertical, but the change was not as marked as in the case of Ba, and it was not parallelled by the roughly proportionate decrease in clockwise perceptions which characterizes Ba's shift of perceptual tendency. The relative weakness of Br's vertical bias is shown by the irregularity of his

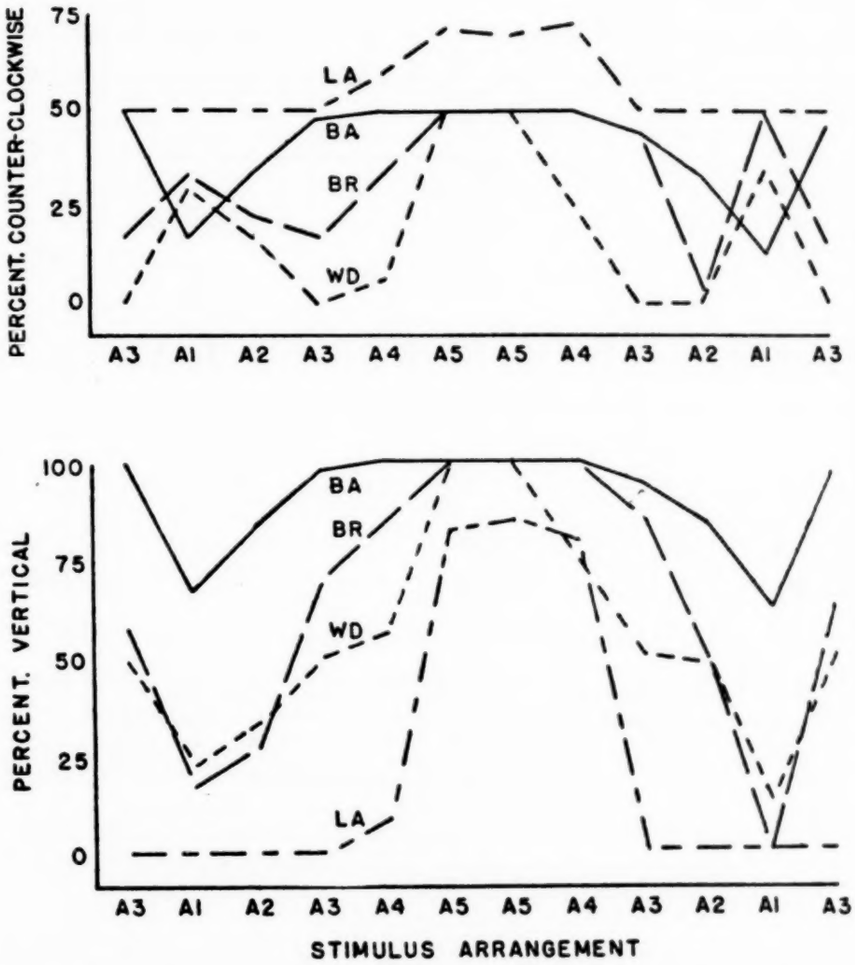


FIG. 4. Mean percentage of vertical and counter-clockwise perceptions reported for each of the stimulus arrangements of Part III. The position of any stimulus arrangement on the abscissa corresponds to its position in the order of presentation.

curve in Fig. 5. Wd remained 'clockwise' for the 'A3' stimulus arrangement. He perceived vertical-counter-clockwise movement with presentation of stimulus arrangements 'A4X' and 'A5X,' and horizontal-counter-clockwise movement with presentation of stimulus arrangements 'A1Y' and 'A2Y.'

These results support the hypothesis advanced in the discussion of the results of Part I, since they provide further evidence of the existence of vertical and other biases varying in intensity in different observers. The vertical biases of Ba and Br appeared for the first time after the factor of shortest distance overcame the clockwise perceptual tendencies of these observers.

Br, Ba and Wd were very much surprised when their perceptions changed from clockwise to counter-clockwise as

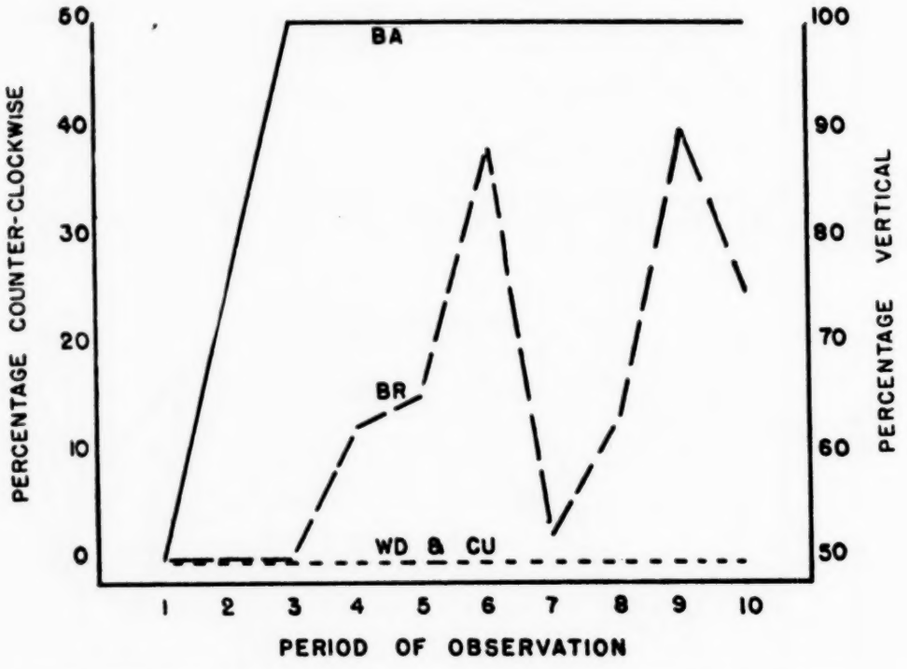


FIG. 5. Percentage of vertical and counter-clockwise perceptions reported for presentations of the 'A3' stimulus arrangement in Part III.

a result of the introduction of the factor of shortest distance in Part III. It is clear from this circumstance that very strong expectation can not always prevent the appearance of a new type of perception when the conditions of stimulation favor it. A perceptual tendency which is overcome by the factor of shortest distance may be regarded as comparatively weak since the factor of shortest distance combined with voluntary efforts to control the direction of movement was found to be incapable of affecting the dominant perceptual biases of some observers.

The influence of the factor of perseveration is shown in Fig. 4. On account of the shift in dominant perceptual tendency of some observers in the early periods of observation of Part III, the mean percentages of Fig. 4 are based on the last six periods of observation of Fig. 3. The effect of the factor of perseveration is evident in the carry-over from the 'A3' stimulus arrangement to the 'A4' and 'A2' arrangements where these immediately follow the 'A3' arrangement. The proportion of vertical perceptions is higher for the 'A4' arrangement when it follows the 'A5' arrangement than when it follows the 'A3' arrangement. Similarly, there are more horizontal perceptions for the 'A2' arrangement when this follows the 'A1' arrangement. This can be seen clearly in the curves of Wd and Br in Fig. 4. Other perseverative effects are apparent elsewhere in the series. Different observers, as Fig. 4 shows, are not affected to the same degree by this factor.

In Fig. 4 the symmetry of the individual curves indicates how well the percentages of a given kind of perception agree for earlier and later presentations of any stimulus arrangement. Allowance must be made, in judging the symmetry of the curves, for the influence of the factor of perseveration.

One observer, La, perceived horizontal movement for presentations of all stimulus arrangements except 'A4' and 'A5.' His perceptions for the 'A4' and 'A5' arrangements showed a definite counter-clockwise bias.

Part IV

Procedure.—In Part III, those subjects whose dominant perceptual tendency was vertical were relatively unaffected by the changes made in the conditions of stimulation. The procedure used in Part III was therefore abridged for these observers and supplemented by a series of stimulus situations designed to test the strength of the vertical bias of these subjects when this was acting in opposition to the factor of perseveration. Four presentations of each of the following stimulus arrangements were given during each period of observation: A3Y, B3Y, C3Y, D3Y, E3Y, F3Y, G3Y, H3Y, J3Y, H3X, G3X, F3X, E3X, D3X, C3X, B3X, A3X.

In this procedure the 'A3' arrangement was rotated in small steps 90 degrees to the left and back again. If the subjects maintained the same direction of rotation as the stimulus was changed from 'A3Y' to 'J3Y,' or from 'H3X' to 'A3X,' the factor of perseveration can be considered to have overcome the vertical bias.

Results.—One 'vertical' subject, Ka, did not perceive rotary movement of any sort when either the 'E3' stimulus arrangement or that on either side of the 'E3' arrangement were presented. The 'E3' arrangement did not give rotary movement for this subject when it was presented in Part II.

The results for three other 'vertical' observers are given graphically in Fig. 6. These results provide further support

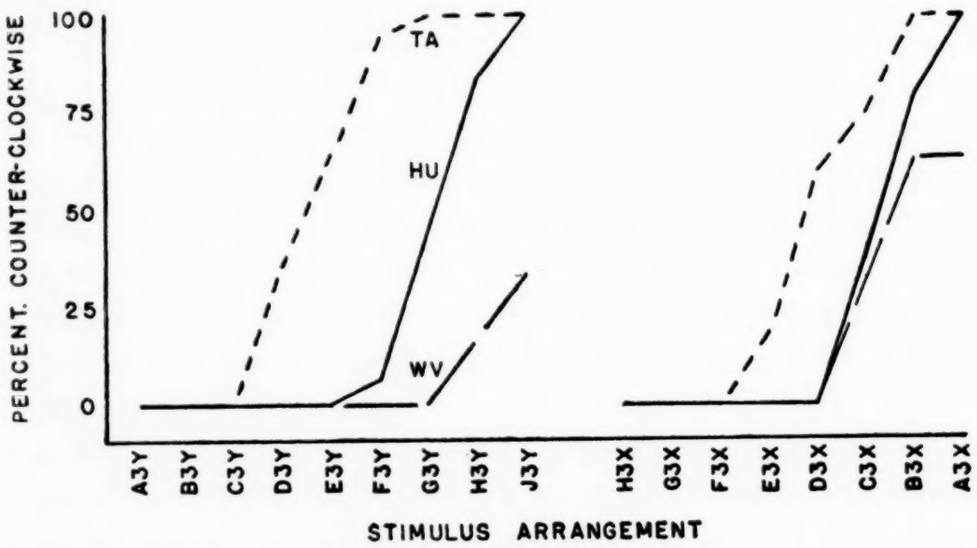


FIG. 6. Mean percentage of counter-clockwise perceptions in Part IV.

for the hypothesis of the existence of different degrees of vertical bias. Both in the 'A3Y' to 'J3Y' series and in the 'H3X' to 'A3X' series, a shift from clockwise to counter-clockwise perception after the 'E3' arrangement was presented indicates that the tendency to see movement as vertical had overcome the factor of perseveration. In both series, Ta shifted from clockwise to counter-clockwise first, Hu and Wv later. The early shift of Ta can not be wholly accounted for, however, on the assumption of a strong vertical bias, since he changed from clockwise to counter-clockwise movement before the 'E3' stimulus arrangement was presented. The vertical component of counter-clockwise movement becomes greater than the horizontal component for the first time in the 'A3Y' to 'J3Y' series when the 'F3Y' arrangement is presented. If the change from clockwise to counter-clockwise movement occurs with presentation of

the 'F3Y' arrangement or later (as is true of Hu and Wv) it can be attributed to the influence of a vertical bias. If the shift from clockwise to counter-clockwise movement occurs before presentation of the 'F3Y' arrangement (as is true of Ta) it must be attributed to some other factor, specifically to the influence of a counter-clockwise bias.

It appears that a vertical (or horizontal) bias may exist in conjunction with a weaker rotational bias which affects perception only when it is not acting against the dominant bias.[3] Ta showed such a subordinate rotational bias both in Part II and Part IV; his preference for counter-clockwise movement is indicated by his shift from clockwise to counter-clockwise movement in advance of the presentation of the 'E3' stimulus arrangement in Part IV. In Part III, Ba manifested a similar subordinate bias for clockwise movement in addition to his final dominant vertical bias. When he shifted from vertical to horizontal perceptions at the 'A1' end of the series, the shift occurred when the 'A1X' stimulus arrangement was presented; in other words, he saw horizontal-counter-clockwise movement much less easily than horizontal-clockwise movement. The perceptions of La in Part III indicate a dominant horizontal bias and a subordinate counter-clockwise bias.

A comparison of the perceptions of Ba and Ta reveals an interesting difference in the relative strength of their respective perceptual tendences. The vertical bias of Ba, although it controlled his perceptions of movement for presentations of the 'A3' stimulus arrangement in Part III, was clearly much less strong than that of Ta, as the much smaller number of horizontal perceptions reported by Ta attests. Further evidence of the relative weakness of the vertical bias of Ba is provided by his perceptions when the stimulus situations of Part IV were presented to him. He saw only clockwise movement. If his results had been included in Fig. 6, they would be represented by a curve coin-

[3] That tendency which determines the perception consistently obtained from the 'A3' arrangement may be arbitrarily termed, for convenience, the dominant tendency. The failure of some observers to perceive movement of any sort when the 'E3' arrangement is presented makes that pattern a less satisfactory standard.

ciding with the baseline. The combined effect of his subordinate clockwise bias and the factor of perseveration may be considered to have overcome his vertical bias, and as a result the shift from clockwise to counter-clockwise movement which the other 'vertical' subjects experienced did not occur.

This analysis indicates the feasibility of assigning intensity values to dominant and subordinate perceptual tendencies, and shows how the hypothesising of such intensity differences is useful in explaining the directional variations of apparent movement reported by the subjects in this experiment.

Part V

Procedure.—In Part V two exposure times were used: 150 sigma, the exposure time used in earlier parts of the experiment, and 500 sigma. The exposure time of the first pair of lights presented was always the same as that of the second pair. The following procedure was designed to determine to what extent the direction of apparent movement could be voluntarily controlled. Four presentations of each of the following stimulus arrangements were given: A1Y, A1X, A2Y, A2X, A3Y, A3X, A4Y, A4X, A5Y, A5X.

In the 'A1' and 'A2' arrangements the factor of shortest distance favored the perception of horizontal movement; in the 'A4' and 'A5' arrangements the factor of shortest distance favored the appearance of vertical movement. The distance between the lights was equal in the 'A3' arrangement.

The entire group of presentations was repeated four times. Before the first two repetitions the observer was instructed to try to make the movement clockwise; before the last two repetitions he was instructed to try to make the movement counter-clockwise. The following table shows the changes in the instructions and conditions of stimulation.

Series	Exposure Time	Direction of Movement Observer Was Instructed to Try to Perceive
1	150 sigma	Clockwise
2	500 "	"
3	150 "	Counter-clockwise
4	500 "	" "

The observer was told when the order of presentation was about to be changed from the 'Y' order to the 'X' order, and vice versa. This made it unnecessary for the observer to attempt to anticipate the changes; a possible source of distraction was thus eliminated.

Results.—The results are shown graphically in Fig. 7. The superiority of voluntary control of movement when the longer exposure time was used is evident. The change from 150 to 500 sigma evoked contradictory reports in regard to the quality of the movement. Most of the observers considered the movement to be better with an exposure time of

500 sigma; one, however, found it so much poorer that he saw almost no movement at all for the first two periods of observation with the longer time. There was also disagreement as to whether the movement was slower or faster with the exposure time of 500 sigma.

The records of two observers, Ta and Hu, are not shown on the graph, as there was no clear superiority of control of

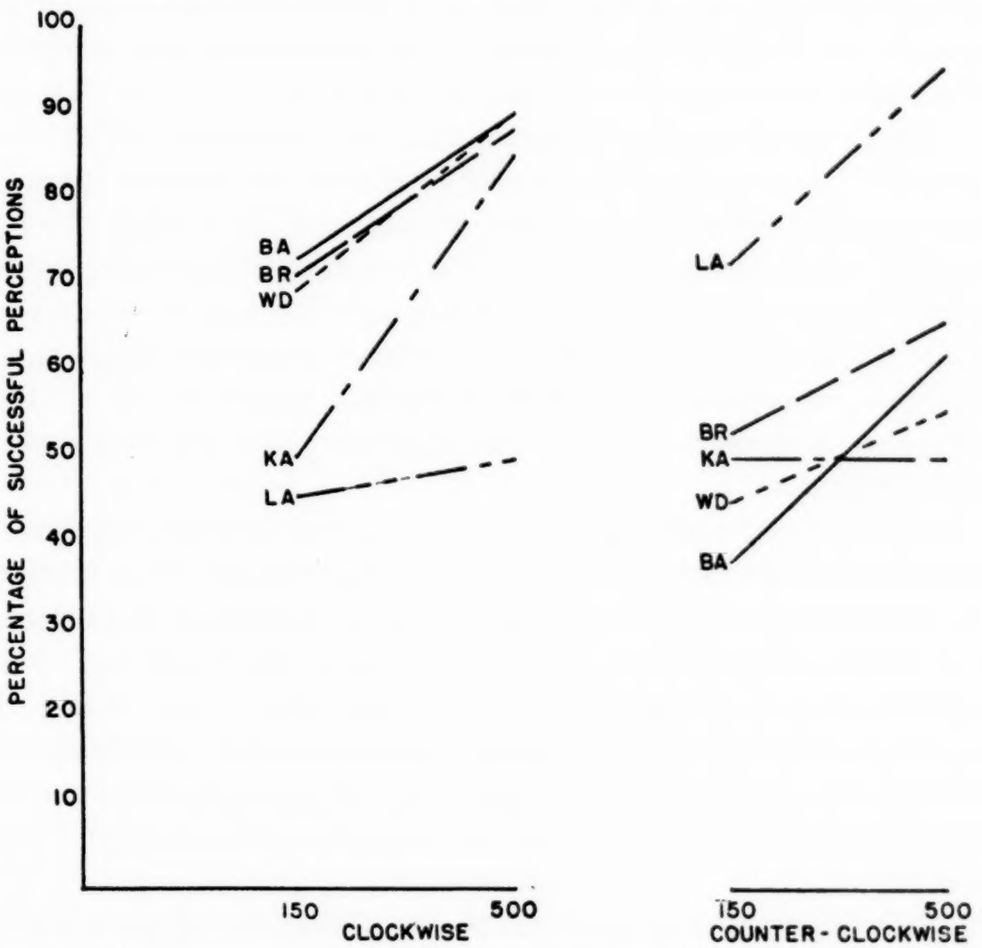


FIG. 7. Mean percentage of successful perceptions (perceptions of the kind of movement the observer was instructed to try to perceive) in Part V.

either kind of rotational movement. The counter-clockwise bias of Ta was not in evidence in this part of the experiment or under the similar conditions of Part III; it was only operative in those series in which the 'A3' stimulus arrangement was rotated to the left.

Four observers, Br, Ka, Wd and Ba, were much more successful when they tried to make the direction of movement

clockwise. This result for three of these subjects, Br, Wd and Ba, was in accordance with the tendency to see movement as clockwise which they had manifested in earlier parts of the experiment. Ka was able to see horizontal movement for the first time when the longer exposure time was used for presentation of the 'A1X' stimulus arrangement. His perceptions provide another illustration of the co-existence of a vertical and clockwise bias of different intensities. Observer La showed an unmistakable preference for counter-clockwise movement, as he did in Part III.

For two periods of observation the fixation point was removed in order to determine whether its absence would have a differential effect upon the perceptions obtained with each of the exposure times. It was found that the removal of the fixation point had no apparent influence on success in control of movement. The observers reported that their fixation invariably wandered when the fixation point was omitted, and that observation was on that account made more difficult.

A variety of methods of controlling the direction of movement was reported. In a similar experiment von Schiller (**5**, 196) found that "a change in the direction of movement was accomplished most successfully, not when the new perception was visualized, but when an effort was made to experience it in terms of motor adjustment." Several observers reported a similar superiority of control when motor imagery was used; others found visual representation more effective.

Voluntary efforts to control the direction of movement are more successful when operating in conjunction with individual perceptual tendencies. A perceptual bias assists an observer in overcoming the influence of the factors of perseveration and shortest distance. When the observers whose records are given in Fig. 7 were trying to perceive that kind of movement for which they had previously exhibited a perceptual bias, the number of successful perceptions was increased by about 25 percent over the number of successful perceptions obtained when their efforts at voluntary

control were acting in opposition to their perceptual tendencies.

Summary

Perceptions of the direction of apparent movement have been found to be characterized by wide individual differences under the conditions of stimulation used in this experiment. This diversity in individual observations indicates the extreme complexity of the interaction of the basic determinants of these perceptions. Lack of knowledge of the physiological processes underlying perceptual activity makes it impossible to explain these changes in perception, except as they can be related to the influence of hypothetical 'biases,' 'factors' and 'tendencies.' With the use of such concepts the results of experimentation on directional variations of apparent movement can be made more intelligible.

Individual biases may include both a vertical or horizontal directional tendency and a rotational (clockwise or counterclockwise) directional tendency. These component directional tendencies generally differ from each other in intensity sufficiently so that one consistently determines the perception when the stimuli are equidistant and so placed as to define a square lying on its side. The relative strength of a perceptual bias can be estimated by the resistance it offers to a change in the perception of movement when the conditions of stimulation are modified to favor such a change.

(Manuscript received December 11, 1935)

References

1. DeSilva, H., An experimental investigation of the determinants of apparent visual movement, *Amer. J. Psychol.* 1926, **37**, 469–501.
2. Hartmann, L., Neue Verschmelzungsprobleme, *Psychol. Forsch.*, 1923, **3**, 319–396.
3. Kelly, E. L., and Miles, W. R., Demonstration unit for patterns of apparent movement, *Amer. J. Psychol.*, 1934, **46**, 475–477.
4. Neuhaus, W., Experimentelle Untersuchung der Scheinbewegung, *Arch. f. d. ges. Psychol.*, 1930, **75**, 315–458.
5. Schiller, P. von, Stroboskopische Alternativversuche, *Psychol. Forsch.*, 1932–33, **17**, 179–214.
6. Waals, H. G. Van der, and Roelofs, C. O., Optische Scheinbewegung, *Zeit. f. Psychol.*, 1930, **114**, 241–288.
7. Wertheimer, M., Experimentelle Studien ueber das Sehen von Bewegung, *Zeit. f. Psychol.*, 1912, **61**, 161–265.

STUDIES IN RETROACTIVE INHIBITION: VII. RETROACTIVE INHIBITION AS A FUNCTION OF THE LENGTH AND FREQUENCY OF PRESENTATION OF THE INTERPOLATED LISTS [1]

BY JOHN A. McGEOCH

Wesleyan University

Introduction

The usual procedure in experiments upon retroactive inhibition, when the interpolated material is of the same kind as the original, has been to interpolate an amount equal to that used as the original learning material. It is in most experiments desirable to maintain this equality of relative amount. In a search for the basic conditions of retroactive inhibition it is, however, important to inquire whether degree of inhibition is a function of the relative lengths of the original and interpolated materials. An answer to this question will, in the first place, extend our knowledge of the variables with which degree of inhibition is correlated and should, in the second, assist a theoretical interpretation of the phenomenon.

The immediate problem of this experiment was to measure the decrements in retention produced by interpolated lists of 2 different lengths and of 3 different frequencies of presentation. One list was equal in length to the original list and one was but half the length of the original. The 3 frequencies of presentation were 4, 8 and 16. The influence of the variable of length was the major problem, but the variations in length introduce, also, variations in degree of learning and the 3 different frequencies have been employed as one means of controlling degree of learning.

[1] This experiment has been assisted by a Grant-in-Aid from the National Research Council, and was done at the University of Missouri.

PROCEDURE

The learning material, both original and interpolated, consisted of lists of two-syllable adjectives. The lists were presented at a 1.8-second rate on an electrically driven drum and were learned by the anticipation method. Each original list contained 16 adjectives plus an initial adjective which served as a cue and was not learned. The original lists were given a constant frequency of 8 presentations. Three of the interpolated lists contained, in addition to the cue-word, 16 adjectives, and 3 contained 8. In order to control by one means the variable of degree of learning, both lengths were given 3 different frequencies. There are, therefore, 6 work conditions in addition to the 2 identical rest conditions which serve as the standard of comparison.

Prior to the beginning of the main experiment each subject was given 2 practice periods in each of which he learned 2 lists to a criterion of 3 perfect trials in succession. At the first practice period the 2 lists contained 8 adjectives each and at the second the 2 lists contained 16 adjectives each. The subjects were, thus, given practice upon 2 samples of each of the lengths which they were to learn in the experiment. They were given, also, 2 relearning presentations of each list at the end of the practice period in order to acquaint them with relearning as well as with learning. The 2 shorter lists were given at the first period in order to avoid discouraging the subjects by the immediate introduction of a difficult task.

Under the rest conditions the subject was given 8 presentations of the original list. Upon the completion of the eighth presentation he turned immediately to reading from a book of humorous short stories of about one page each under instruction to select the best one from those which could be read in the time allowed. After approximately 18 minutes of reading he was asked to record his judgment.[2] At the end of exactly 20 minutes from the end of the eighth learning presentation the subject began to relearn the original list or, if it had not been learned to a criterion during the 8 repetitions, to continue learning. The performance after the interval will be referred to throughout as relearning. Relearning was to a criterion of 3 perfect repetitions in succession. Two identical rest conditions have been employed in order to make the basic control records more reliable.

Under each of the work conditions the subject was given 8 presentations of the original list followed, after the 30 seconds required to reset the drum for the presentation of another list and to instruct the subject that a second list was to be learned, by either 4, 8 or 16 presentations of the second or interpolated list. The remainder of the twenty-minute period was filled by reading humorous short stories as under the rest condition. At the end of the interval the original list was relearned to a criterion of 3 successive perfect trials and the interpolated list was given 2 relearning repetitions. These 2 repetitions were for the purpose of maintaining relative constancy of attitude on the part of the subject toward original and interpolated lists. Had there been no measure of the retention of the interpolated list, he might have gained the impression

[2] The subjects considered the reading of the stories and the selection of the best one as a second part of the experiment and there is good reason to believe from their comments and other behavior that there was virtually no rehearsal during the interval. Some means must be taken to prevent rehearsal on the part of untrained subjects and the reading of material of this sort has proved to be satisfactory for this purpose. The 'rest' period is, of course, one of rest only as defined by this particular activity and insofar as it involves no formal learning of the kind required by the remainder of the experiment.

that it was less important than the original list and might therefore have practiced it under later conditions with less effort and without intent to retain.

The frequencies and the general plan of the experiment are summarized in Table I.

TABLE I

OUTLINE OF THE EXPERIMENTAL CONDITIONS

Condition	Original Learning	Interpolation During 20-min. Interval	Relearn
Rest.....	Learn 16_8	Reading	1
Rest.....	Learn 16_8	Reading	1
Work 1....	Learn 16_8	Learn 16_4 Reading	1, 2
Work 2....	Learn 16_8	Learn 16_8 Reading	1, 2
Work 3....	Learn 16_8	Learn 16_{16} Reading	1, 2
Work 4....	Learn 16_8	Learn 8_4 Reading	1, 2
Work 5....	Learn 16_8	Learn 8_8 Reading	1, 2
Work 6....	Learn 16_8	Learn 8_{16} Reading	1, 2

The major figure represents the length of the list and the subscript represents the frequency of presentation. The figures 1 and 2 under "Relearn" refer to the original and interpolated lists respectively. The original list was relearned to 3 perfect repetitions in succession and the interpolated list was given 2 relearning presentations.

The 8 conditions are the same except for the introduction under the work conditions of the learning of a second list. It can be seen from the table that the 2 lists have equivalent frequencies. Equal frequencies applied to lists of different lengths do not mean equality in number of items presented or in the degree of learning of these items. Equality of total number of items presented during the interpolated learning obtains, however, between Work 1 and Work 5 and between Work 2 and Work 6. In the first pair a total of 64 items is presented and in the second a total of 128. The extent to which these equal numbers of total items have meant equality of degree of learning can be measured in one respect by the numbers of correct anticipations, and the results will presently be discussed.

The conditions were gone through in a counterbalanced practice order which gave equal average position to each condition. The lists were distributed over the conditions in a way to control variations in difficulty. Forty subjects, all of them college students of sophomore rank or higher, went through each of the 8 conditions of the experiment. Of this number 23 were men and 17 were women. Sex differences in the learning of these materials are so small as to be insignificant for the present problem.

Since there are 2 equivalent rest conditions, N under rest should be 80. In one group of 16 subjects, however, a few unavoidable disturbances occurred during one rest condition with some of the subjects which seemed clearly to introduce a variation in conditions. Rather than run the risk of loading the data by discarding only the few subjects who seemed to be affected, the records of this condition have been discarded for the whole group. This can be done without destroying the effectiveness of the counterbalancing. N is, therefore, 64 under rest.

The 40 subjects were taken through the experimental conditions at the rate of one condition per day. They were divided into 3 groups, each of which had a different experimenter. The groups contained 16, 16 and 8 subjects, respectively. Each subject learned individually, of course. The term 'group' means only the subjects who served for any one experimenter. Each experimenter had had practice in at least one previous experiment upon retroactive inhibition and with the anticipation method

of learning. The 3 were given careful training upon the method of this experiment and their work was supervised by the writer throughout. The data obtained by each experimenter have been treated separately, but the 3 sets of data show an agreement which is well within the limits of deviation ascribable to the small samplings and show no variations which seem attributable to the factor of the experimenter. The 3 sets have, therefore, been combined in the tables which follow.

Retention has been measured in terms of 4 basic scores: (1) Recall 1, which is the number of correct anticipations at the first presentation after the interval; (2) Recall 2, which is the number of correct anticipations at the second relearning presentation; (3) Relearning 1, which is the number of trials required to relearn up to but not including the first perfect repetition; and (4) Relearning 2, which is the number of trials required to relearn up to but not including the criterion of 3 perfect trials in succession.

Results

Relative Degrees of Original and Interpolated Learning.—It is desirable to check the equal frequencies of repetition of the original lists by some more direct measure of learning, and the total number of correct anticipations made during the 8 repetitions is one measure for this purpose. The mean number of correct anticipations per subject at each condition, together with the sigma of the mean, is shown in Table II.

TABLE II

Mean Correct Anticipations per Subject

(N is 64 under the rest condition and 40 under all work conditions)

Lists	Condition						
	Rest	Work 1	Work 2	Work 3	Work 4	Work 5	Work 6
Original M	49.35	43.90	51.67	48.42	51.65	49.12	47.57
σM	1.71	2.25	2.78	2.39	2.34	2.59	2.49
Interpolated M		14.02	51.72	152.30	12.15	36.14	99.55
σM		1.12	2.64	5.75	.65	2.27	1.73

The means of the original lists vary from condition to condition but no difference is 3 times its sigma. When it is remembered, moreover, that the differences are in terms of single items and that the data are from 8 learning trials, the differences become of small significance. They exist, however, and must be taken into account in interpreting the results.

The relative degrees of learning of the interpolated lists may be measured in terms of the numbers of correct anticipations. Such a measure is necessary if the inhibitory

effect of these lists is to be interpreted. It can be seen from Table II that with both lengths of lists the mean number of correct anticipations increases as the frequency. The means do not, however, have the same significance for the 2 lengths. The means under Work 1 and Work 4 are, for example, nearly of the same size, but in Work 1 the correct anticipations have been distributed over 3 repetitions [3] of a sixteen-item list while in Work 4 they have been distributed over 3 repetitions of an eight-item list.

It is necessary to know what the distributions are in the lists of different lengths. The correct anticipations may have been concentrated at the first 2 or 3 positions in each case; they may have been distributed over the whole list; or they may have been variously concentrated. If the appearance of an item above the threshold of anticipation were unequivocal, it would be possible to use the frequency of supraliminal items as a measure of the distribution. An item may, however, be anticipated on an early trial but not at later ones and an adequate criterion of appearance above the limen is, for the present purpose, lacking. The mean number of correct anticipations per subject at each serial position gives the briefest and, as it seems, the best picture of distribution, and has been used (Table III).

The serial position values speak for themselves and only a few representative comparisons will be mentioned. Under Work 1 there are only 5 positions which have a mean of one item or more. At the intermediate positions only a part of the group successfully anticipated a single item. From Work 1 to Work 3 there is, as would be expected, a steady increase in the means at each position. Under Work 2, where the frequency of presentation of sixteen-item lists was the same as that of the original lists, the serial position values are closely similar to those of the original lists. The composite values of the latter are given for purposes of comparison. There is a similar increase at each position from Work 4 to Work 6.

[3] The list was given 4 repetitions but at the first there is no possibility of a correct anticipation.

TABLE III

MEAN NUMBERS OF ANTICIPATIONS AT EACH SERIAL POSITION

Serial Position	All Original Lists (16_8)	Interpolated Lists Condition					
		Work 1 (16_4)	Work 2 (16_8)	Work 3 (16_{16})	Work 4 (8_4)	Work 5 (8_8)	Work 6 (8_{16})
1	5.60	1.83	5.47	12.40	2.33	5.83	13.88
2	4.92	1.55	4.70	11.05	1.73	5.45	12.93
3	4.29	1.20	3.87	10.00	1.73	3.98	12.25
4	3.42	.85	4.25	9.68	1.53	4.68	12.33
5	2.93	.55	3.33	7.90	.93	3.40	11.90
6	2.22	.50	1.75	8.75	1.35	3.55	11.70
7	2.06	.52	1.82	7.25	1.20	4.15	11.65
8	2.40	.45	1.87	8.33	1.35	5.10	12.90
9	2.20	.38	2.30	7.53			
10	1.98	.70	2.72	9.45			
11	2.48	.68	2.35	9.48			
12	2.45	1.40	2.73	8.65			
13	2.10	.42	2.82	9.03			
14	2.22	.75	2.88	9.33			
15	2.96	.88	3.75	11.00			
16	4.65	1.38	5.10	12.40			

The serial position values for the 2 lengths cannot be compared directly. The means under Work 4 are, in general, larger than those under Work 1, which affords an equivalent number of presentations, and are smaller than those at Work 2 and Work 3. The means toward the ends of the list under Work 5 are of the same order as those towards the ends under Work 2, but the values for the intermediate positions are larger under Work 5. Much the same is true when Work 6 is compared with Work 3. The usual serial position curves are approximated in each of the 7 columns.

There is, here, a situation in which (1) every interpolated list save one (Work 2) differs from the original list in numbers of correct anticipations at corresponding serial positions, (2) there are wide differences between the means at corresponding positions of interpolated lists of equal length but of differing frequencies of presentation, and (3) there are differences between the 2 different lengths of lists with respect to the absolute serial position values. It is of great interest, then, to inquire what degrees of retroactive inhibition result from these varying relations between original and interpolated

lists and between interpolated lists of differing length and of differing serial position values.

Measures of Retention and of Retroactive Inhibition.—The means and the sigmas of the means of the 4 basic measures of retention under each of the experimental conditions are given in Table IV. It is necessary to examine first the secon-

TABLE IV

RECALL AND RELEARNING SCORES

Condition	Interpolation	Recall 1	Recall 2	Relearning 1	Relearning 2
Rest.....		7.95±.46	10.81±.40	9.85± .87	14.12±1.31
Work 1.....	16_4	4.77±.56	8.45±.56	11.00±1.10	16.12±1.77
Work 2.....	16_8	4.95±.61	7.87±.62	10.55±1.05	15.62±1.63
Work 3.....	16_{16}	3.75±.65	7.37±.60	10.57±1.01	15.77±1.71
Work 4.....	8_4	6.52±.71	10.27±.53	8.82± .88	14.60±1.50
Work 5.....	8_8	5.25±.71	8.75±.57	8.77± .99	15.00±1.54
Work 6.....	8_{16}	5.15±.71	8.77±.56	8.90± .77	14.05±1.34

dary problem of the influence upon retention of the variable frequencies of presentation of the interpolated lists, since a knowledge of this influence will be required in interpreting the results upon relative length. For the purposes of the secondary problem the 2 lengths of interpolation will be treated separately. A lower score under work than under rest means, of course, a decrement from the interpolated learning, *i.e.* retroactive inhibition.

In terms of Recalls 1 and 2 every work condition produces some retroactive inhibition. The differences between the scores under rest and under each of the work conditions are 3 times the sigma of the difference in every case except that of Work 4.[4] No difference between the mean trials required to relearn to either 1 or 3 perfect trials under rest and any work condition is 3 times its sigma,[5] although the differences show, for each length separately, some consistency in the direction of the difference.

[4] The differences and their sigmas are, for Recall 1, as follows: Work 1, 3.18 ± .72; Work 2, 3.00 ± .76; Work 3, 4.20 ± .80; Work 4, 1.43 ± .85; Work 5, 2.70 ± .84; and Work 6, 2.80 ± .84. The corresponding data for Recall 2 are: 2.36 ± .69; 2.94 ± .74; 3.44 ± .72; .54 ± .66; 2.06 ± .70; and 2.04 ± .69.

[5] In every case the sigma of the difference is larger than the difference and there is nothing to be gained by reproducing them here.

No difference between any 2 recall or relearning scores within the 2 groups of 3 work conditions each is as much as twice its sigma. When the interpolated list is 16 items long, 4 and 8 repetitions yield nearly equal means in terms of Recall 1, while 16 repetitions yield a considerably smaller mean.[6] There is in terms of Recall 2 a small but consistent decline as frequency increases. Both relearning scores show a slightly larger number of trials required to relearn when the frequency of the interpolated list is 4 than when it is 8 or 16. The most probable conclusion is, however, that the 3 different frequencies of interpolation of sixteen-item lists produce equal amounts of inhibition in terms of relearning, while in terms of recall they yield amounts of inhibition which have a slight tendency to vary directly as the frequency.

When the interpolated list is 8 items long, Recall 1 decreases uniformly as frequency increases, while Recall 2 decreases as frequency increases from 4 to 8 but is equal for 8 and 16 repetitions. Neither relearning score exhibits any regularity of tendency or any differences of considerable size. The conclusion here, as in the case of the sixteen-item list, is that recall scores have a slight tendency to decrease as interpolated frequency increases while relearning scores do not vary with frequency.[7]

The influence of the factor of relative length upon the inhibitory effect of the interpolated lists, which is the major problem of the experiment, may be seen more readily if the differences between the measures of retention under rest and

[6] It is possible that the smaller mean number of correct repetitions at original learning under Work 1 has contributed to the production of a smaller mean recall at the first relearning trial.

[7] The tendencies of the recall scores of this experiment are in agreement with those of an earlier experiment, while those of the relearning scores are not uniformly in agreement. (*Cf.* J. A. McGeoch, The influence of degree of interpolated learning upon retroactive inhibition. *Amer. J. Psychol.*, 1932, 44, 695–708.) The former experiment employed lists of nonsense syllables as both original and interpolated materials and compared interpolated frequencies of 6, 11, 16, 21 and 26 after an original frequency of 11. Since materials, lengths and frequencies differ in the two experiments, it is not surprising that differences appear in the results. In the present experiment, which was not designed primarily to study the influence of frequency, the variations in frequency were introduced as aids to a comparison of the inhibitory effects of different lengths of list.

work are expressed in the percentages which these differences are of the score under rest. This can be done with relearning scores in the same way as with recall scores since the frequency at original learning is a constant. The results in terms of the 4 basic scores are given in Table V. These data

TABLE V

PERCENTAGES OF RETROACTIVE INHIBITION AND OF FACILITATION

Condition	Interpolation	Recall 1	Recall 2	Relearning 1	Relearning 2
Work 1.....	16_4	40.0	21.8	11.6	14.1
Work 2.....	16_8	37.7	27.2	07.1	10.6
Work 3.....	16_{16}	52.8	31.8	07.3	11.7
Work 4.....	8_4	17.9	05.0	+10.5	03.4
Work 5.....	8_8	33.9	19.0	+11.0	06.2
Work 6.....	8_{16}	35.2	18.9	+09.6	+00.5

are derived from the means of Table IV and must be interpreted in the light of the raw data and of their statistical reliability. Examination of the table will show that, in terms of every one of the 4 measures of retention, *every interpolation of 16 words produces a greater decrement in retention than does any interpolation of 8 words*, regardless of relative frequency and of attendant degree of learning. The differences between the means of retention scores after the 2 different lengths of interpolation, *i.e.* between the values for work which are represented in these percentages, are in no case as much as 3 times sigma. The consistency of the direction of the differences may, however, be taken as a guarantee of the validity of the direction.

This uniformity of direction appears in spite of wide differences in interpolated frequency and degree of learning. Frequencies of 8 and 16 applied to an eight-adjective list yield a much greater total number of correct anticipations than a frequency of 4 applied to a sixteen-adjective list; and 16 repetitions of the shorter list yield more correct anticipations than 8 repetitions of the longer list. Nevertheless, the longer list produces the greater inhibition.

If, moreover, we consider the distribution of the correct anticipations over the lists, the preponderant influence of

length or amount becomes still more apparent. Four repetitions of the sixteen-item interpolated list yield a very small number of correct anticipations at any position and a mean of at least one anticipation at only 5 positions, yet a relatively large amount of inhibition results and more than is produced by any eight-item list, although at each of the serial positions in the latter, and especially under Work 5 and Work 6, there are relatively large numbers of correct anticipations. Analogous comparison may be made between each pair of conditions. Number of items is clearly a determining condition of degree of inhibition over and above degree of learning as here measured.

Differences in frequency of interpolation within a single length of list are, as we have seen, accompanied by regular variations in retroaction only when retention is measured by recall scores. The extent of these variations is not great and the smallness of their extent, together with the absence of regular variation in terms of relearning, would lead one to expect relatively small differences in retroaction from differences in the degree of learning of different lengths. In comparisons of amounts of inhibition resulting from different lengths of interpolated list, the variable of degree of learning must, however, be controlled. It has been necessary, therefore, to vary frequency and to present the results of such variation in order to make sure that the variations in inhibition which are correlated with the differences in relative length are not produced wholly or in part by differences in degree of learning. The latter is no unitary condition which can be expressed by a single measure. It has many dimensions and is measurable in many different ways. The two measures which have been used in the present experiment are representative. Comparisons of the degrees of learning of lists of different lengths involve differences in proactive and retroactive inhibition within the individual lists and in other aspects of intra-list organization. The measures used are only partial measures of these factors. They are, however, unequivocal measures as far as they go. Inhibition has shown itself to be a function of length of list independently

of wide variations in these measures of degree of learning and, within the limits of the experimental conditions and method of measurement, the conclusion that the longer list produces the greater inhibition is valid.

It is, however, important to note that the differences between the inhibitory potencies of the 2 lengths of the interpolated lists, while it is consistent, is not always large. The eight-item lists uniformly produce inhibition in terms of both recall scores, and the amounts under work conditions 5 and 6 are not greatly less than those produced by the sixteen-item lists in work conditions 1 and 2. In terms of recall scores an interpolated list half the length of the original list exerts a considerable inhibitory effect. In terms of relearning scores, the shorter list has, however, practically no inhibitory influence, but shows instead some facilitative effect.

Inhibition as a Function of Serial Position.—It at once becomes necessary to ask whether, especially at recall, inhibition varies differently with serial position under the 2 amounts of interpolated material. Is inhibition, for example, concentrated at 8 serial positions, when only 8 items are interpolated, or is it distributed over all positions?

The percentage of inhibition has been computed for each serial position in the original list under each condition. This has been done, in terms of the 2 recall scores, by subtracting the mean number of items correctly recalled at each position under work from the mean number recalled at the corresponding position under rest and then by dividing the remainder by the number under rest. A similar procedure has been used with the number of trials taken to relearn the item at each position to the point beyond which no further errors were made.

It would be cumbersome to reproduce the percentages of inhibition at each of the 16 serial positions. Their tendencies at recall are faithfully represented by the percentages of Table VI, in which the serial positions are grouped in intervals of 4.[8] Work conditions 1 to 3, it will be remembered, have

[8] The inhibition in each group of 4 serial positions has been found by adding together the recall scores of each group and performing with the means of these records the operations already described.

TABLE VI

PERCENTAGES OF INHIBITION AT SUCCESSIVE SERIAL POSITIONS IN GROUPS OF FOUR

Condition	Recall 1 Groups of Serial Positions				Recall 2 Groups of Serial Positions			
	1–4	5–8	9–12	13–16	1–4	5–8	9–12	13–16
Work 1.....	43.9	46.5	29.3	34.6	21.4	39.7	21.8	05.8
Work 2.....	43.9	37.2	26.8	42.3	22.6	34.9	23.6	30.4
Work 3.....	57.8	62.8	41.5	48.1	25.0	44.3	29.1	30.4
Work 4.....	25.0	02.3	17.1	23.1	13.1	09.5	−03.6	07.2
Work 5.....	48.4	23.3	29.3	30.8	23.9	17.5	20.0	15.9
Work 6.....	50.0	27.9	31.7	25.0	19.1	30.2	20.0	11.6

interpolated lists 16 items in length, and work conditions 4 to 6 have interpolated lists 8 items in length.

The percentages of inhibition, when computed for groups of serial positions, show the same general tendencies as do those computed for total lists. With exceptions, which are not important for the present purpose, 16 interpolated items produce more inhibition than do 8, and within each length of interpolation inhibition tends to vary directly as frequency.

The major question is whether with an interpolation of 8 items inhibition shows a different distribution over the serial positions than it does with an interpolation of 16. The answer is that there is no clear evidence that it does. With an interpolation of 8 items there is some inhibition in terms of Recall 1 at every group of positions in the original list. There is, moreover, no differential distribution of amount. Inhibition at the last 4 positions is less than at the first 4, but the same is true for each of the interpolations of 16 items. It is greater at positions 9–12 than at positions 5–8, thereby reversing the relation under interpolation of 16 items, but the differences between the 2 groups of positions with an interpolation of 8 items are seldom large. It is doubtful if the relatively greater reduction of inhibition at positions 5–8 in the eight-item list should be taken to mean a valid differential distribution.

When we turn to Recall 2, we find, likewise, some inhibition at every group of serial positions except one (condition 4, positions 9–12). Some of the relationships already mentioned

in discussing Recall 1 are reversed, a fact which warns against taking the relations in either case too seriously. There is no evidence for any marked change in the intra-serial pattern of inhibition with a shift from 16 to 8 interpolated items.

When it is remembered that these percentages of inhibition at groups of serial positions are computed from relatively small numbers of items at each position and, therefore, from unstable values, even considerable differences between positions must be regarded skeptically. The outstanding and consistent result is that inhibition is distributed over all groups of positions regardless of the length of the interpolated list. The lack of a consistent tendency in Recalls 1 and 2 toward any differential distribution of amount prevents any conclusion that a difference appears.[9]

The percentages of inhibition at relearning have been found in the same way. They are uniformly small and they show more definitely than do those for recall scores a complete absence of any differential distribution of inhibition over the 16 serial positions.

Discussion

Inhibition has been found to be a function of the relative amounts of material in the original and interpolated lists. The factor of relative amount has more influence under these conditions than does relative frequency or relative number of correct anticipations per item. Inhibition has, moreover, been found at all serial positions of the original list, whether the interpolated list contains 8 or 16 items, and there is no clear differential distribution of inhibition over the serial positions as a function of change in amount of the interpolated material.

These results appear in spite of great differences between the numbers of correct anticipations in the original and interpolated lists. The extent of the differences in degree of learning between original and interpolated lists and from condition to condition may be seen by consulting Tables II and III.

[9] The data at each individual serial position were carefully studied before the positions were grouped. The statements made for the grouped positions are supported by the examination of the separate components.

The major facts which require interpretation are (1) that inhibition is a function of relative length of list regardless of relative frequency when frequency is varied over a considerable range and (2) that inhibition fails to show any clear differential distribution over the serial positions of the original list with a halving of the length of the interpolated list and with great relative differences in the numbers of correct anticipations at the corresponding serial positions of the 2 lists.

The perseveration theory of Müller and Pilzecker [10] fails to throw light upon these results. It might, of course, be argued that the learning of 16 interpolated items means a more intense interpolated activity than the learning of 8 items and hence a greater blocking of the perseveration of the neural correlates of the original list. The fact that the eight-item lists also produce inhibition and that this inhibition is distributed over all of the 16 positions of the original list could be interpreted in terms of a general blocking of perseveration by a somewhat less intense activity than that of the longer interpolated list. It is very doubtful, however, whether a low frequency of the longer list is as intense an activity as a much higher frequency of the shorter list. To make the perseveration theory applicable, one must endow it with all of the properties implied by the results and in the absence of knowledge about perseveration this endowment leads at once to circularity. The best that can be said for the theory is that it is indeterminate with respect to the present results.

The transfer theory, as it is usually stated, provides no immediately ready interpretation, but it seems to be the more reasonable one with which to attempt an interpretation. It has been supported repeatedly by experiments which have found inhibition to be a function of the relationships between the original and interpolated materials. The variable in the present experiment is the amount of the interpolated material, which places the results immediately within the scope of the

[10] Müller, G. E., and Pilzecker, A., Experimentelle Beiträge zur Lehre vom Gedächtniss, *Zsch. f. Psychol.*, 1900, Erg. Bd. 1, 1–300.

theory. The transfer theory is, moreover, susceptible to more direct experimental test. We shall inquire, therefore, in more detail how the transfer theory could account for the results.

The greater inhibitory potency of the longer list was, in the first place, to have been expected on the basis of results of McKinney and McGeoch.[11] They found considerable amounts of overt transfer of items from one list to another, with transfer from interpolated list to original much greater than from original to interpolated. The sixteen-item interpolated list contains more items to carry over to the recall of the original list and to become confused with the members of that list. Other things being equal, therefore, it should be more inhibitory than the eight-item list, and this has been found to be the case.

The amounts of inhibition from the interpolated learning of 8 items could not have been predicted on the grounds of the McKinney and McGeoch experiment. On those grounds alone the eight-item list should have one-half as much inhibitory influence as has the sixteen-item list. If we consider recall scores only and compare Work 4 with Work 1, 2 or 3, the eight-item list has less than half as much, but if we compare Work 5 and 6 with Work 1, 2 or 3, it has more than half as much. These comparisons are, of course, complicated by differences in degree of learning as well as in length, and are not to be taken too seriously. It is useless to devote time to them until we know more about the interrelations of the determiners of inhibition. Reasoning from the numerical possibilities of transfer alone, the results in terms of relative amounts produced by the 2 different lengths follow only partially from those of the earlier experiment. The fact that the greater inhibition from 16 items does follow from these earlier results, however, points clearly to the transfer theory.

A more detailed interpretation of the facts requires, however, that we know how the decreased inhibition from the shorter interpolated list, together with the distribution of this inhibition over all serial positions, could have resulted from

[11] McKinney, F., and McGeoch, J. A., The character and extent of transfer in retroactive inhibition: disparate serial lists. *Amer. J. Psychol.*, 1935, 47, 409–423.

transfer between the 2 lists. The 2 lists are, under each work condition, similar to the extent that the words are all two-syllable adjectives, learned by the same procedure, and this similarity is almost certainly one determiner of the obtained inhibition. It does not of itself account for the results peculiar to this experiment. It is necessary to inquire what other factors are common to the 2 lists and the way in which any common factors may have operated to bring about the results. Three of the possible common factors will be described.

(1) The *experimental context* is clearly the same for the 2 lists. We may include in the context the exposure apparatus, the experimenter, the room and its furniture, the instructions and the resulting set of the subject, and the operations by which the subject learns the lists. Each of these aspects of the context is a separate common factor and some of them may be analyzed into several sub-factors. We may, however, classify them for the present purpose under the general heading of experimental context. This context is present, of course, in every work condition of this kind. Let the context be connected by the subject with each list and to some extent with each item in each list, and let the context be designated by A, the original and interpolated lists by B and C, respectively. In any work condition, A is first connected with B and then with C. At the learning of C there may be, by virtue of associative inhibition, some intrusion of items from B, and C, once learned, may be expected to produce some reproductive inhibition which amounts to a disruption or weakening of the organization of B.

The obtained retroactive inhibition is, then, insofar as this interpretation is valid, conceptually reduced to the reproductive inhibition paradigm with experimental context as the common term. To state it differently, the specificity of the intra-serial connections of B is disrupted and at the recall of B the common factor A may arouse items from either B or C. This interference or confusion is what we are measuring by the experimental operations of the retroactive inhibition experiment. It will be greater with a sixteen-item interpolated list than with one of 8 items because the common

A is connected with an equally large number of items in original and interpolated lists under the first condition, and the possibilities of confusion are numerically greater. There is no reason to suppose that the common context is specific to any part of the list, so that the similar distributions of inhibition over the serial positions of the 2 different interpolated lengths is a reasonable expectation. This still does not account for the fact that the eight-item list produces in most comparisons more than half as much inhibition as the sixteen-item list. This is possibly better accounted for by the influence of a second common factor.

(2) This second factor is *areal position in the series.* Subjects, when they are learning serial lists by the anticipation method, think of each item not only in connection with other items in the list but also as being located roughly at some position in the list. Let the following rough designations stand for the ways in which a subject may think of the serial positions of the items, regardless of the length of the list: at the beginning, near the beginning, just before the middle, about at the middle, just past the middle, near the end, and at the end. The subject may not use these words but he thinks of the item in ways which are at least analogous and which these will represent and he learns them in this schematic framework. Each position or area may become connected with one or more items in the list; in lists of 16 items each one, and especially those designating intermediate positions, may be connected with several items. The connections may be of varying strengths and their strengths may change in unknown ways with continued repetition of the list, but it is fair to assume that they exist, and that each word or class will be connected with more items in a longer list than in a shorter.

The classifications 'at the beginning,' 'near the middle,' etc., are not specific to any list but may be used by the subject, with variations, perhaps, in the learning of any list. In the learning of an original and an interpolated list, each classification becomes connected with 2 different sets of items. When, therefore, the subject comes to the relearning of the original

list, he has 2 sets of items connected with the general meaning of 'at the beginning' or 'near the middle' and this common meaning may operate to elicit either or both sets with consequent blocking or with incorrect recall. There is no reason to suppose that the influence of this common factor will be twice as great when the original and interpolated lists are both 16 items long as when one is 16 and the other 8. It might well be nearly as great in the second case as in the first, although the relatively greater specificity of the class-item associations in the shorter list, and the consequently small probabilities of confusion would lead to the expectation of at least a somewhat smaller amount of inhibition. The inhibition should, as has been found, be distributed over the entire original list.

(3) A third possible factor is *the connection of items in both original and interpolated lists with a common item in the language organization of the subject.* All of the words in all of the lists were two-syllable adjectives. All of them have widely ramifying connections in the language organization of the subjects. It is easily possible that many of the items in any given pair of lists should be associated with some common item in the language organization of the subjects. Let Item A in List 1 arouse Item X in the subject's language organization (but not in either list). Let Item K in List 2 also arouse Item X. Let corresponding connections be made between other words in the 2 lists and third items which are common to the pairs of words, and let these common items enter into the organization of the list by the subject. Any given instigating word at the recall of the original list may then arouse a common Item X (or meaning) connected with an item in each list. This will lead to interference and confusion. The common items will have no necessary association with serial position and we would expect to find some inhibition at each position from each interpolation.

This can fit the specific results of this experiment only if we expand the range of the common factor. Let Item A in an eight-item interpolated list sometimes be connected with Item X which is also connected with Items E and H in the

original list. We can then readily imagine the occurrence of the results. This multiple connection with a common factor should be less frequent than single connections. Hence the longer list should produce the more inhibition, but not necessarily twice as much, and inhibition from either interpolation should be distributed over all serial positions.

These three common factors do not exhaust the possibilities. They illustrate, however, ways in which the conditions which are characteristic of the retroactive inhibition experiment may be analyzed. The results of such analysis lead to the hypothesis that retroactive inhibition reduces to the conceptual paradigm of reproductive inhibition. Any factor whatever, so long as it is common to the original and interpolated lists, may function as the common term in this paradigm.[12] The usually employed dimensions of similarity offer clear examples of such a common term. We need, for an understanding of the decrements following interpolated learning, a detailed analysis of all possible common factors and of the conditions of their operation. The three possibilities which have been described are few among many. They have been described because they are especially called for by the present results.[13]

Summary

Original lists of 16 two-syllable adjectives have been given a standard frequency of 8 presentations and have been relearned after 20 minutes. The interval was filled either by reading humorous short stories (rest) or by the learning of a second or interpolated list of adjectives (work). In 3 work conditions the interpolated lists contained 16 items and were given frequencies of 4, 8 or 16 repetitions. In 3 other work

[12] *Cf.* McGeoch, J. A., and McGeoch, G. O., Studies in retroactive inhibition: VI. The influence of the relative serial positions of interpolated synonyms. This Journal, 1936, **19**, 1–23.

[13] Nothing has been said thus far about the fact that relearning scores show either facilitation or very slight inhibition when 8 items are interpolated. This is apparently an aspect of the more general problem of the relations between different measures of retention and of their differential susceptibility to inhibition. It requires discussion in the light of more data than are now at hand.

conditions the interpolated lists contained 8 items and were given frequencies of 4, 8 or 16.

1. Every interpolation of 16 adjectives produces a greater decrement in retention than does any interpolation of 8 adjectives, regardless of relative frequency and of attendant degree of learning. Inhibition is, thus, under these conditions, a function of the relative lengths of the original and interpolated lists, while relative frequency is comparatively unimportant.

2. There is no differential distribution of inhibition over the serial positions of the original list as a result of change in the amount of the interpolated materials. The inhibitory influence of the shorter list is, thus, not specific to any serial positions.

3. These results can more readily be interpreted by the transfer theory than by the perseveration theory. The hypothesis is stated that the transfer takes place according to the paradigm of reproductive inhibition and that three of the most important common factors are experimental context, areal position in the series, and connection with a common item in the language organization of the subject.

(Manuscript received December 12, 1936)

THE EFFECT OF PRACTICE ON POSITIVE TIME-ORDER ERRORS

BY HERBERT WOODROW AND LELAND H. STOTT

University of Illionois

The purpose of the investigation was to determine the effect of practice upon time-order errors, in the comparison of successive stimuli, when at the beginning of practice the time-order errors for all variables were positive. By practice is here meant simply the prolongation, through repeated sittings, of the series of comparisons.

The desirability of such a study was indicated by results previously obtained [1] in which the time-order errors for all variables were initially *negative*. In this previous investigation, the standard was an empty interval of 1000 σ duration, and the variables ranged from 20 percent shorter to 20 percent longer than the standard. In the case of all variables, the subjects' judgments at first indicated that the variable seemed longer when given second, *i.e.*, following the standard, than when given first, *i.e.*, before the standard. These initial negative time-order errors, however, showed a marked change with practice. The change was uniformly in the positive direction. As a result of this change, the large negative errors shown initially by the variables longer than the standard became much smaller negative errors, while the small negative errors originally manifested by the variables shorter than the standard soon changed into positive errors.

In the attempt to understand these peculiar practice effects, the negative errors initially manifested were assumed to be due to the influence of the indifference-interval, an interval presumably shorter than any of the intervals used in the experimental series. It was stated that this phenomenon of an initial negative time-order error for all variables could

[1] Woodrow, H., The effect of practice upon time-order errors in the comparison of temporal intervals. *Psychol. Review*, 1935, **42**, 127–152.

be regarded as one which would result if during the pause of 2 secs between the two compared intervals, the first interval 'gravitated' downwards, towards the indifference-interval. The shift towards positivity which occurred in the errors with practice may then be explained on the assumption that as the subject becomes more habituated to the experimental series, the influence of the indifference-interval is for the most part obliterated, and replaced by that of the average length of the intervals presented in the entire preceding series of comparisons, *i.e.*, by what is termed the 'series-average.' [2]

By assuming that similar relations would obtain when the indifference-interval was *longer* than any of the compared variables, a prediction could be made of the results to be expected in the case here studied. All variables could be expected, initially, to show positive time-order errors. The change to be expected as the result of practice, as the influence of the indifference-interval was supplanted, or partially so, by that of the series-average, would depend upon the length of the variable relative to the length of the standard, or series-average. With variables shorter than the standard, the original positive time-order errors should persist or become enhanced; with variables longer than the standard, the original positive errors should grow smaller or even change to negative ones; and with the zero-difference variable, the original positive errors should change to smaller positive ones or completely vanish.

The expectations outlined above, on the basis of general principles, have all been realized, as will be shown by the data to follow. They have been realized as regards the initial time-order errors (positive for all variables) and as regards the *direction* of the expected change with practice in the time-order errors. Not having formulated the governing principles as quantitative laws, the *amount* of change in the errors produced by practice could not be foreseen. It was therefore not known whether to expect the initially positive time-order

[2] For the concept of the series-average as a remote standard see Woodrow, H., Weight-discrimination with a varying standard, *Amer. J. Psychol.*, 1933, **45**, p. 409; also, Lauenstein, O., Ansatz zu einer physiologischen Theorie des Vergleichs und des Zeitfehler, *Psychol. Forsch.*, 1933, **17**, p. 174.

errors of the longer variables to become merely less positive or to change to negative ones. Likewise it could not be foreseen whether the initially positive errors with variables shorter than the standard would become greatly enhanced or only slightly affected. It must be confessed, however, that the results were not, when viewed quantitatively, entirely in accord with the expectations based on previous work. The discrepancy lay in the fact that while the changes in the time-order errors with practice were, in general, in the predicted direction, they were not large enough to be statistically significant except in the case of the increase in the positive time-order errors of the shorter variables.

Time-order errors do not tell the whole story. A time-order error, measured in terms of its effect upon percentages of variable longer reports, is a difference between two percentages: (1) the percentage of reports signifying variable longer with the order *V–S*; and (2) with the order *S–V*. The time-order error varies only with the relative magnitude of these two percentages; but each of the percentages, taken separately, may show interesting changes with little or no change in the time-order error. It follows that a thorough analysis of the effects of practice can be made only by studying separately for each time-order the changes in the percentages. Through such study, the conclusion has been reached that, under the conditions of the present study, while the influence of the series-average may have increased considerably with practice, the indifference-interval, instead of being supplanted by the series-average, stubbornly continued to exert its effect in but very slightly diminished degree. This conclusion stands in considerable contrast with one previously reached in the case where the series-average was *longer* than the indifference-interval, instead of shorter, as here, to the effect that the influence of the indifference-interval was rapidly and to a marked degree supplanted by that of the series-average. The difference between these two cases is explained on the assumption that it is much more difficult to adapt to lengths shorter than the indifference-interval than to lengths longer than the indifference-interval.

The main experiment here reported was carried out with a group of 43 subjects, to whom were presented for comparison at each of 10 sittings a series of 150 pairs of tonal durations. The procedure followed the method of constant stimuli, in which a standard having a duration of 500 σ was compared with five variables differing from the standard by − 20, − 10, 0, + 10, and + 20%. The effect of practice has been studied by comparing the results for the last two sittings with the first two sittings. The data for any one sitting have not been fractionated, that is, successive portions have not been separately analyzed, because in the present study the sequence of the comparison pairs was such that it was impossible to divide the sitting into several equal segments in which the sequence was identical. It should be remembered, therefore, that the errors calculated from the first sitting taken as a whole may differ considerably from those which would have been revealed by the very first reports. A supplementary experiment was conducted with a second group, composed of 35 subjects. The work with this section was merely a repetition of the first two sittings with the first section.

The tones used as stimuli and the apparatus which produced them, which was rather complex, has been described elsewhere[3] in detail. The tones, though rather impure, had a pitch of 1000 dv, and were sounded by a loud speaker connected with an audio-oscillator. The durations were controlled by making and interrupting the electrical circuit by means of a synchronous motor used as a rotary contact-device. All durations and intervals were carefully checked by measurements made by means of a Hipp's chronocope, in turn controlled by a Klopsteg fall-apparatus. Special arrangements prevented the occurrence of 'clicks' at the onset or cessation of the tones. The flash of a miniature electric lamp, suspended above the loud-speaker, gave the 'ready' signal, 1.5 secs before the onset of the first tone of each comparison-pair.

In investigations of this sort it is imperative that the results permit a calculation of the time-order for each variable. Such results can be obtained only by presenting each variable before the standard in one half of the trials and after the standard in the other half. Such a procedure was accordingly followed. Of course, with zero difference, there is no distinction between variable and standard. Arbitrarily, in this case, the first given interval will be here termed the variable. The number of trials with the zero-difference variable (trials in which the standard was compared with itself) was equal to the total number of trials (including the two orders *V–S* and *S–V*) with any other variable.

The subjects recorded their judgments in three categories, longer, shorter, and equal. In the tables here presented the equal judgments have been divided between longer and shorter. Table I shows the percentage of reports, after this division, signifying that the variable was longer (in the case of the zero-difference variable, that the first duration was longer). Since a subject's reports always referred to the second interval, a report signifying variable longer in the case of the order V_1 would be recorded on the subject's record blank by a minus sign, indicating that the standard was the shorter. By comparing the percentages in the two columns given for each variable, V_1 and V_2, it may be conveniently observed how much oftener the variable was indicated to be longer than the standard when it came first than when it came second. The difference in the two frequencies gives the time-order error for that variable. Thus, according to Table I, when the 20% shorter variable came first it seemed longer than the standard 16.2% of the trials; whereas when it followed the standard it seemed longer in only 4.5% of the trials. The time-order error is a function of the difference

[3] Stott, L. H., Time-order errors in the discrimination of tonal durations, J. EXPER. PSYCHOL., 1935, **18**, pp. 741–766.

between these two percentages—what function, precisely, has never been ascertained. It will suffice for the present purposes to use this difference as an index of the magnitude of the time-order error. Whenever the percentage of variable longer reports is greater with V_1 than with V_2, the time-order error is, by convention, termed positive, and when the percentage is greater with V_2 than with V_1, negative.

For convenience, the time-order errors, which may be calculated from the data of Table I, have been tabulated in Table II.

TABLE I

THE EFFECT OF ORDER OF PRESENTATION UPON THE PERCENTAGES OF REPORTS SIGNIFYING VARIABLE LONGER

The line labelled "Variable" indicates by what percentage the variable differed in length from the standard.

Each cell contains a percentage based on 645 comparisons, in the case of Section I, and on 525 comparisons, in the case of Section II.

Total no. of comparisons represented in the table is 37,300.

Variable	−20%		−10%		0%		+10%		+20%	
Order	V_1	V_2	V_1	V_2	V_1	V_2	V_1	V_2	V_1	V_2
Section I										
1st day	16.2	4.5	27.0	15.4	56.1	43.9	75.2	65.8	85.1	81.9
2d day	20.3	4.1	41.4	16.1	57.7	42.3	72.8	63.1	82.6	81.1
9th day	22.6	4.9	43.7	19.7	54.0	46.0	73.2	70.4	83.5	81.4
10th day	28.4	6.4	40.8	17.3	54.1	45.9	75.1	68.8	81.7	86.4
Section II										
1st day	24.1	6.4	35.2	19.2	55.4	44.6	71.5	61.6	83.0	75.9
2d day	31.3	10.3	46.1	21.2	56.6	43.4	67.6	60.4	74.0	73.8

Table II shows that, with a pause of 1.5 secs, the time-order errors for all variables were positive at the first sitting. It follows that there must be an indifference-duration at a length greater than that of the longest variable, that is, beyond 600 σ. This result is in harmony with Stott's finding that, with a pause of 1.5 secs, the indifference-duration has a length of about 900 σ.[4]

If the average time-order error for the first two sittings be calculated from the data for Sections I and II, with equal weight given to each section, the percentage time-order errors for the variables, in order from the − 20% to the + 20% variable, are as follows: + 16.7%, + 19.4%, + 12.9%, + 8.8%,

[4] Stott, L. H., *op. cit.*

TABLE II

TIME-ORDER ERRORS IN TERMS OF THE DIFFERENCE IN THE PERCENTAGE OF "VARIABLE LONGER" REPORTS DUE TO DIFFERENCE IN ORDER

Variable		−20%	−10%	0%	+10%	+20%
Section I.....	1st day	+11.7	+11.6	+12.2	+9.4	+3.2
	2d day	+16.2	+25.3	+15.4	+9.7	+1.5
	9th day	+17.7	+24.0	+ 8.0	+2.8	+2.1
	10th (last) day	+22.0	+23.5	+ 8.2	+6.3	−4.7
Section II.....	1st day	+17.7	+16.0	+10.8	+9.9	+7.1
	2d (last) day	+21.0	+24.9	+13.2	+7.2	+0.2
	Change in Time-Order Error					
Section I.....	1st to last day	+10.3	+11.9	− 4.0	−3.1	−7.9
	Standard dev.	*1.7*	*2.3*	*2.7*	*2.5*	*2.3*
	Critical ratio	6.1	5.1	1.5	1.2	3.4
Section I.....	1st to 2d day	+ 4.5	+13.7	+ 3.2	+0.3	−1.7
II.....	1st to 2d day	+ 4.3	+ 8.9	+ 2.4	−2.7	−6.9

and + 3.0%. These differences, which are differences between two percentages each calculated on the basis of 2340 trials, are all statistically significant (from 6 to 17 times the σ_{diff}) except possibly that obtained with the + 20% variable, which is only 2.5 times the σ_{diff}. In general the initial time-order errors increase in magnitude with increase in the distance of the variable from the indifference-duration. The one exception to this trend is that the error is smaller in the case of the − 20% variable than in that of the − 10% variable; but to make the various percentage time-order errors strictly comparable, the differences represented by them would need to be subjected to an absolute scaling technique. It will suffice, here, to point out that the error + 16.7% in the case of the − 20% variable is the difference between 23.0 and 6.3%; whereas the error of + 19.4% in the case of the − 20% variable is that between 37.4 and 18.0%. The latter difference is almost certainly the smaller in any absolute sense.

The effect of practice on the time-order errors is shown in Table II by the difference between the last and first day, Section I, and also by that between the second and first day, both sections. The general trend may be stated as follows:

(1) the error with variables shorter than the standard becomes more positive; (2) with variables longer than the standard, less positive, or negative; and (3) with the zero-difference variable, less positive. Several exceptions to this generalization appear if the first sitting is compared with the second one. In this case the zero-difference variable shows an exception in the case of both Sections I and II, in that the positive error is slightly greater on the second day. In the case of Section I, the error for the + 10% variable also was more positive on the second day than on the first. Some of the changes are not great enough to be regarded as significant. A rough index of significance is the ratio of the change to its standard deviation, when the latter is taken as the standard deviation of the difference between two differences, each of which is in turn a time-order error, that is, a difference between two percentages. Standard deviations of the change with practice in time-order errors and the ratios of the changes to their standard deviations are given in Table II for the changes which occurred in the case of Section I between the first and last sitting. Of these changes, those with the zero-difference and the + 10% variable are not significant.

Before attempting to account for the changes with practice on the part of the time-order errors, it will be helpful to push the analysis further and study the changes occurring in the percentages for each order taken separately. These changes, calculated as differences between the 1st and last days' percentages, are given in Table III.

Both the heading given Table III and the plus and minus signs appearing in the table are based on theoretical assumptions concerning the explanation of time-order errors. It has been assumed that the time-order error is due to a discrepancy between the true objective length of the first of a pair of stimulus-durations and the length with which (as indicated by the subjects' judgments) the subject appears to be comparing the second stimulus-duration. This latter length, termed the effective length of the first stimulus-duration, is presumably somehow represented by the internal condition of the subject at the time of occurrence of the second stimulus-duration. An example will clarify the

TABLE III

CHANGE WITH PRACTICE IN THE EFFECTIVE LENGTH OF THE FIRST STIMULUS-DURATION

Change in the effective length of the first interval is expressed in terms of the change in the percentage of trials that the first interval was indicated to be longer.

Variable	−20%		−10%		0%	+10%		+20%	
Order	V_1	V_2	V_1	V_2	V_1	V_1	V_2	V_1	V_2
Section I									
1st to last day	+12.2	−1.9	+13.8	−1.9	−2.0	−0.1	−3.0	−3.4	−4.5
Standard dev.	*2.3*	*1.3*	*2.6*	*2.1*	*2.8*	*2.4*	*2.6*	*2.1*	*2.3*
Critical ratio	5.3	1.5	5.3	.9	.7	0.0	1.2	1.6	2.0
1st to 2d day	+ 6.7	−1.4	+ 8.1	−2.7	−2.8	+0.2	−5.1	−1.3	−2.4
Section II									
1st to 2d day	+ 7.2	−3.9	+10.9	−2.0	+1.2	−3.9	−1.2	−9.0	+2.5
Standard dev.	*2.8*	*1.8*	*3.1*	*2.5*	*3.1*	*2.8*	*3.0*	*2.5*	*2.7*
Critical ratio	2.6	2.1	3.5	.8	.4	1.4	.4	3.6	.9

procedure. With Section I, the − 20% variable was at the first sitting indicated to be longer than the standard, when it came first (column V_1, Table I) in 16.2% of the trials. On the 10th day, however, it was indicated to be longer in 28.4% of the trials. Obviously, then, the subjects' reports changed with practice as they would have been expected to change had the − 20% variable when given first grown longer. This fact is expressed in Table III by recording as + 12.2 the change with practice in the effective length of the first duration, in the case of the − 20% variable, order V_1. Each value in Table III thus shows the change with practice in the case of a particular time-order in the percentage of judgments indicating that the first-given duration was the longer. A plus sign indicates that the change in this percentage was one which would be expected providing the first-given duration had increased in length (without change in any other determining factor); and, similarly, a minus sign indicates that the change in percentage was one to be expected had the sole change in conditions been a decrease in the length of the first-given duration.

The general trends, shown by Table III, may be very simply stated: (1) the effective length of the first duration, when it consisted of a variable shorter than the standard,

became longer with practice; and (2) in all other cases, the effective length of the first duration, irrespective of whether it was the standard or the variable, grew shorter. While the preceding generalization is true for the results of Table III taken as a whole, one or more exceptions appear in the case of each section. Moreover, such a general statement conceals some interesting differences in the effect of practice on the different comparison-pairs and different orders. One interesting fact is that most of the effect of practice was exerted on the time-order V_1, and, with this time-order, on variables shorter than the standard rather than on variables longer than the standard. In fact, in the case of Section I, there was no significant change with this order except when the variables were shorter than the standard. In the case of Section II, however, there did appear a significant change in the case of the + 20% variable (V_1). With the standard first (V_2), probably none of the changes observed in the case of either Section I or II should be regarded as significant.

It is possible that the changes in the percentages, which have here been interpreted as indicative of a change in the effective length of the first-given stimulus, are a resultant of changes in this effective length and changes due to improved acuity or keenness of discrimination. There is, however, no evidence of any improvement in acuity. The most reliable, if indeed not the only reliable, change produced by practice was directly the opposite of that which would result from an increase in acuity alone. This change, as already indicated, consisted in a pronounced increase in the percentage of variable longer judgments when the variables were given first and were shorter than the standard. From these considerations, it appears very doubtful that the practice given, which was uncorrected practice, exerted any significant effect upon acuity.

The small magnitude of the changes occurring with practice in the case of standard first or in that of a variable longer than the standard first was somewhat surprising. In the case of empty intervals, all of which were longer than the indifference-interval, the standard, when first, appeared rather definitely to increase in effective length. In that study,

the influence of the *shorter* indifference-interval appeared to be greatly weakened by practice and to be supplanted by that of the series-average. In the present study, on the contrary, the indifference-duration, now *longer* than any of the compared durations, appears to have maintained its initial influence throughout with but very little weakening. Of course the two investigations differed in other respects. Continuous durations may behave differently from empty intervals; but data obtained by Stott on the effect of practice in the case of continuous durations longer than the indifference-duration (900 σ when the pause is 1.5 secs) were very similar to those obtained by Woodrow with empty intervals longer than the indifference-interval (presumably 600 σ with a pause of 2 secs). It seems probable, therefore, that the effect of practice depends upon whether the indifference-interval is longer or shorter than the compared stimuli. When the compared intervals are shorter than the indifference-interval, the effect of the indifference-interval persists in much more stubborn fashion than when the compared intervals are longer than the indifference-interval. This conclusion is interpreted to mean that a subject can more readily adapt or attune himself to the standard or series-average length when this is longer than the indifference-interval than when it is shorter. The indifference-interval thus appears as a limit, a *lower* limit, in the temporal lengths to which the nervous system of the subject can respond in an entirely adequate manner.

If it be assumed that, because of the great difficulty of adapting to a duration shorter than the indifference-interval, the effect of this interval is not appreciably diminished with practice, it is possible to understand the results here obtained. The persistence of this remote standard should not interfere with the development of the series-average as a second remote standard. But if each first-given duration continues, as initially, to gravitate [5] upwards, towards the indifference-interval, then it may be said that the series-average itself continues to gravitate upwards. After practice, then, the

[5] For the sense in which this term is used, see Woodrow, H., *Psychol. Rev.*, 1935, 42, p. 141.

first-given duration of each comparison-pair would gravitate towards both the series-average and the indifference-duration; but since the series-average itself would be gravitating upwards towards the indifference-duration, the level towards which the first-given durations were gravitating might not be appreciably lower than initially. Even the variables longer than the standard might continue to gravitate upwards, and almost as much as at the beginning. The standard, when given first, while with practice tending to gravitate towards itself, would continue to gravitate upwards, and the errors would continue positive. Thus, practice would produce little change in any of these cases, that is, with standard first or a variable longer than the standard first. But with variables shorter than the standard, the tendency to gravitate upwards towards the series-average would be enhanced, due to the formation of the series-average as an effective remote standard,—and all the more enhanced because the series-average was itself deflected in the direction of the still longer indifference-duration. It is thus possible to account for the increase of 12 or 13% (Section I) in the variable longer judgments when the variable was given first and was shorter than the standard.

Summary

Practice in the comparison of short tonal durations was given a group of 43 subjects for 10 sittings, each sitting comprising 150 comparisons arranged in accordance with the method of constant stimuli. The standard was a duration of 500 σ and the five variables differed from the standard by -20, -10, 0, $+10$, and $+20$%; so that all the compared durations were considerably shorter than the duration 900 σ, previously determined as the indifference-duration for tones of the pitch and loudness here used and for the same pause, namely, 1½ secs, between compared tones. As was to be expected from the fact that all the variables were shorter than the indifference-duration, initially they all showed positive time-order errors, that is, were judged longer more often as the first than as the second member of a comparison-pair.

With practice (in the sense of mere increase in number of

sittings), the time-order errors became less positive in the case of all variables. Since a time-order error is a difference in the percentage of variable longer judgments due to change in the order from *V–S* to *S–V*, the effect of practice on each of these orders was studied separately. Such study revealed the following changes, occurring as a result of practice, in the percentage of trials that the first duration of a comparison pair was indicated to be the longer: When the first duration was a variable shorter than the standard, there was a marked, reliable increase in the percentage; in all other cases, that is, when the first given duration was either the standard or a variable longer than the standard, there occurred, with a few unreliable exceptions, a slight decrease, a decrease which with one exception was too small to be regarded as significant.

It is pointed out that all the facts noted may be explained on the hypothesis of two remote standards, described in a previous paper. In the present instance, however, it appears necessary to assume that the first remote standard, corresponding to the temporal indifference-duration, represents a physiological limit in the sense that it is impossible to become adequately adapted or attuned to shorter durations. As a result, the second remote standard, corresponding to the average duration of the compared intervals, gravitates towards, or is assimilated by, the first remote standard. All durations constituting the first member of a comparison-pair are regarded as gravitating towards both remote standards: initially, towards only the indifference-duration but after practice towards both the indifference-duration and the series-average. Since the series-average is considered itself to gravitate upwards, towards the indifference-duration, the only marked effect of practice is to increase the effective length of the first-given durations when the latter are shorter than the series-average. The standard and intervals longer than the standard continue to gravitate upwards, towards the indifference-duration, almost as much after practice as initially, even though after practice they may be regarded as pulled slightly towards the series-average as well as towards the indifference-duration.

(Manuscript received November 25, 1935)

THE EFFECT OF INTERPOLATED TIME INTERVALS UPON THE CONTRAST EFFECTS

BY MALCOLM G. PRESTON

University of Pennsylvania

Previous studies by Turner,[1] Arons and Irwin,[2] and Preston [3,4] have been concerned with the fact, first noticed by Fernberger [5] that in psychophysical experimentation with the method of constant stimuli, the subject tends to avoid the repetitive use of the same category of judgment. Turner [6] was inclined to attribute the occurrence of this fact solely to stimulus conditions, but Arons and Irwin [7] and Preston [8,9] have shown this explanation to be inadequate by reason of the fact that contrast effects exist when stimulus conditions are constant, for all categories of judgment (including Equal), and between judgments not temporally proximate in the series. Preston [10] suggests that the contrast effects in the usual psychophysical experiment may be explained by assuming first, that the subject maintains an attitude against the repeating of judgments, and second, that the effective

[1] Turner, W. D., Intra-serial effects with lifted weights, *Amer. J. Psychol.*, 1931, **43**, 1–25.

[2] Arons, L., and Irwin, F. W., Equal weights and psychophysical judgments, J. Exper. Psychol., 1932, **15**, 733–51.

[3] Preston, M. G., Contrast effects and the psychophysical judgments, *Amer. J. Psychol.*, 1936, **48**, 389–402.

[4] Preston, M. G., Contrast effects and the psychometric functions, *Amer. J. Psychol.*, 1936, **48**, 625–631.

[5] Fernberger, S. W., Interdependence of judgments within the series for the method of constant stimuli, J. Exper. Psychol., 1920, **3**, 126–50.

[6] Turner, W. D., *op. cit.*, p. 16.

[7] Arons, L., and Irwin, F. W., *op. cit.*

[8] Preston, M. G., Contrast effects and the psychophysical judgments, *Amer. J. Psychol.*, 1936, **48**, pp. 399–400.

[9] Preston, M. G., Contrast effects and the psychometric functions, *Amer. J. Psychol.*, 1936, **48**, 625–631.

[10] Preston, M. G., Contrast effects and the psychophysical judgments, *Amer. J. Psychol.*, 1936, **48**, pp. 400–402.

value of the standard weight is influenced by the effective value of the preceding comparison weight.

The disposition against repetition has been studied in connection with other kinds of data by Telford,[11] Thorndike [12] and Dodge.[13]

Thorndike's experiment consisted of having a subject give numbers, picked at random, to words which were read by the experimenter. Examination of his data disclosed the fact that fewer cases were found where the subject gave the same number to successive words than would be expected by chance. He also found increasing avoidance from sitting to sitting and from day to day. In explanation of this fact Thorndike believed the subjects to have been influenced by factors extrinsic to the process of thinking and writing, on the grounds that the introduction of an interval of five seconds did not diminish the avoidance of repetition sufficiently to permit the interpretation of the phenomenon on the basis of refractory phase.

While in general Dodge was in agreement with Thorndike's interpretation he attached considerable potential significance to the concept of refractory phase and suggested a crucial experiment to test its applicability. Telford in his 1931 study attempted to relate the concept to the kind of behavior studied by Thorndike by means of the technique suggested by Dodge.

Essentially the method used in this experiment was Thorndike's, although, following Dodge's suggestion, intervals of time of varying lengths were interposed between the reading of the syllables (which Telford substituted for the words used by Thorndike). In general it was observed that the longer the period (up to 16 seconds) the greater the frequency of repeated numbers. On the basis of this observation, Telford wrote: [14]

[11] Telford, C. W., The Refractory Phase of Voluntary and Associative Responses, J. Exper. Psychol., 1931, 14, 1–36.

[12] Thorndike, E. L., The Refractory Period in Associative Processes, *Psychol. Rev.*, 1927, 34, 234–36.

[13] Dodge, R., A note on Professor Thorndike's Experiment, *Psychol. Rev.*, 1927, 34, 237–240.

[14] Telford, C. W., *op. cit.*, pp. 26 and 33.

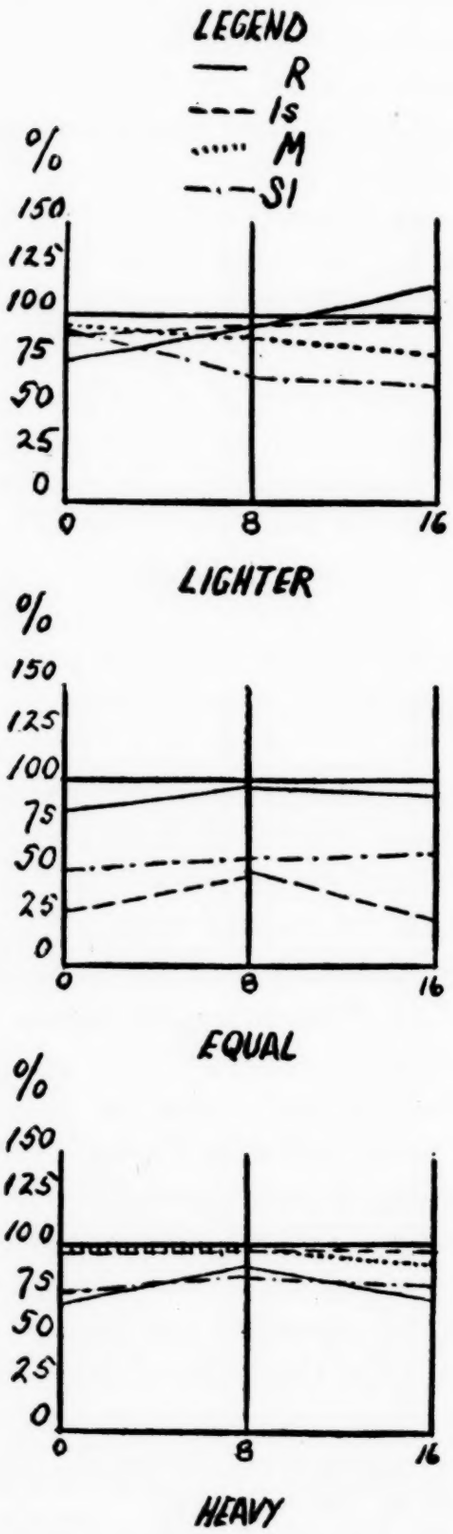


CHART A. Showing percent of expected frequencies of repetition observed in series A, D, and E, for categories lighter, equal and heavier.

"Something very similar to a refractory period seems to be operative here. . . . It seems probable that the effect of a refractory period, followed by a period of hyper-excitability and a subsequent return to normal in the systems involved, is worth considering seriously as a large factor in the explanation of these phenomena. We do not wish to imply that we consider this to be an ultimate explanation. . . ."

The present experiment proposes to investigate the effect of interpolating various intervals of time upon the observed and expected frequencies of judgment sequence in order that the temporal properties of this set may be studied with respect to Dodge's criterion.

Method of Procedure

The experimental and statistical methods used in this experiment are described in detail by Arons and Irwin [15] and Preston.[16] Fourteen pairs of weights, each weighing 100 grams, were presented to the subjects, referred to hereafter as Is, M, R, and Sl, at a rate of 92 beats per minute. During Series A, the subjects lifted weights without interruption between the comparison weight of one pair and the standard weight of the next. During Series D, the subject was instructed to permit 8 beats of the electrically driven metronome to elapse between the lifting of the comparison weight of one pair and the standard weight of the next. During Series E, sixteen beats of the metronome were permitted to elapse between the lifting of each pair. The subjects were instructed to report whether the second weight of each pair was lighter than, equal to, or heavier than the first. The most probable number of each of the nine possible sequences of judgment categories was computed and compared with the observed number. On the basis of this comparison two separate values were obtained, both for sequences of proximate judgments and for sequences with one judgment interposed between the members. First the percent of expected frequencies which were observed to have occurred was computed. Next the value Diff./σ freq. giving the reliability of the *difference* between expected and observed frequencies was computed. The present experimental data are composed of 14,728 judgments equally divided between Series D and E, with Sl, R and M giving 1,872 judgments on each series and Is giving 1248 on each series, and 8748 judgments in Series A, complete details of which have been given in another place.[17] Subject M's data differ from those of the others because he gave no equal judgments under the conditions of Series D and E. Experimentation was begun in May 1933, was concluded in June 1934, and was carried out in the Psychological Laboratory at the University of Pennsylvania.

Results

The results of this work are shown in Tables I, II, III, IV, V, and VI. Tables I, II, and III give the percent of expected

[15] Arons, L., and Irwin, F. W., *op. cit.*, pp. 740–741.

[16] Preston, M. G., *op. cit.*, pp. 391–392.

[17] Preston, M. G., *op. cit.*, pp. 393–398.

TABLE I

SHOWING PERCENT OF EXPECTED FREQUENCIES OF JUDGMENT SEQUENCE OBSERVED WHEN NO INTERRUPTION OCCURS BETWEEN PAIRS OF STIMULI—SERIES A

Subject...	Is		M		Sl		R	
Judgment Sequences	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed
LL....	89	85	95	95	94	108	76	125
LE....	96	128	78	78	84	86	104	93
LG....	107	104	103	103	112	109	97	102
EL....	104	97	56	56	116	113	106	105
EE....	30	38	33	33	51	100	83	95
EG....	109	112	119	114	131	97	124	107
GL....	104	104	107	107	89	90	93	86
GE....	116	97	111	111	140	104	124	110
GG....	95	99	96	96	74	100	66	88

TABLE II

SHOWING PERCENT OF EXPECTED FREQUENCIES OF JUDGMENT SEQUENCES OBSERVED WHEN 8 BEATS ARE INTERPOSED BETWEEN STIMULUS PAIRS—SERIES D

Subject...	Is		M		Sl		R	
Judgment Sequences	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed
LL....	95	93	89	84	66	113	89	122
LE....	118	116	—	—	102	100	104	103
LG....	92	100	104	106	110	97	95	89
EL....	98	114	—	—	135	112	106	100
EE....	47	42	—	—	56	96	96	98
EG....	114	106	—	—	121	100	106	104
GL....	103	102	107	108	85	83	91	105
GE....	102	106	—	—	131	105	109	105
GG....	97	97	98	97	83	101	89	93

TABLE III

SHOWING PERCENT OF EXPECTED FREQUENCIES OF JUDGMENT SEQUENCE OBSERVED WHEN 16 BEATS ARE INTERPOSED BETWEEN STIMULUS PAIRS—SERIES E

Subject...	Is		M		Sl		R	
Judgment Sequences	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed
LL....	98	101	78	102	63	102	115	160
LE....	94	94	—	—	91	104	92	92
LG....	103	100	113	98	121	96	107	91
EL....	121	98	—	—	132	98	96	87
EE....	25	33	—	—	60	104	94	101
EG....	98	111	—	—	120	97	113	104
GL....	99	101	112	100	88	101	98	100
GE....	118	117	—	—	137	96	114	100
GG....	97	96	93	100	75	104	71	99

TABLE IV

SHOWING VALUES OF DIFF./σ FREQ. ON JUDGMENT SEQUENCES—SERIES A

Subject...	Is		M		Sl		R	
Judgment Sequences	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed
LL....	−1.18	−1.67	− .80	− .37	− .34	.51	− .90	.95
LE....	− .23	1.72	− .67	− .34	−1.55	−1.37	.40	− .62
LG....	1.11	.58	.67	.06	1.38	1.02	− .08	.19
EL....	.26	− .23	−1.34	− .67	1.63	1.31	.62	.49
EE....	−2.55	−2.27	−1.16	−1.16	−8.24	− .04	−4.37	−1.42
EG....	.82	1.06	.68	.14	6.13	− .67	5.13	1.53
GL....	.66	.58	1.53	.92	−1.29	−1.20	− .48	−1.02
GE....	1.42	− .14	.41	− .14	7.97	.80	5.07	2.12
GG....	−1.29	− .38	−1.28	− .48	−6.17	0	−5.62	−1.94

TABLE V

SHOWING VALUES OF DIFF./σ FREQ. ON JUDGMENT SEQUENCES WHERE 8 BEATS ELAPSED BETWEEN THE COMPARISON OF EACH PAIR OF WEIGHTS—SERIES D

Subject...	Is		M		Sl		R	
Judgment Sequences	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed
LL....	− .45	− .83	−1.37	−1.84	−2.34	.92	− .34	.67
LE....	1.28	1.14	—	—	.21	0	.42	.19
LG....	.42	.05	.76	1.09	1.22	− .37	− .34	− .80
EL....	− .17	.99	—	—	.36	1.24	.52	− .05
EE....	−2.34	−2.57	—	—	−6.89	− .69	−1.38	− .80
EG....	1.39	.59	—	—	3.95	.05	1.22	.81
GL....	.56	.41	1.29	1.44	−1.83	−2.12	− .65	.28
GE....	.25	.59	—	—	5.96	.95	1.81	1.04
GG....	− .55	− .62	− .83	− .98	−4.04	.28	−1.63	−1.11

TABLE VI

SHOWING VALUES OF DIFF./σ FREQ. ON JUDGMENT SEQUENCES WHERE 16 BEATS ELAPSED BETWEEN THE COMPARISON OF EACH PAIR OF WEIGHTS—SERIES E

Subject...	Is		M		Sl		R	
Judgment Sequences	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed	Sequences Proximate	One Judgment Interposed
LL....	− .32	.08	−2.99	.34	−2.83	.16	1.08	2.75
LE....	− .42	− .44	—	—	− .96	.48	− .91	− .91
LG....	.52	.09	2.47	−.33	2.63	.45	.55	− .69
EL....	1.52	− .15	—	—	3.54	−.29	− .43	−1.49
EE....	−2.59	−2.31	—	—	−9.53	.71	−1.79	.39
EG....	− .16	1.02	—	—	3.76	−.53	2.67	.73
GL....	− .20	.16	2.46	−.05	−1.49	.07	− .14	0
GE....	1.41	1.28	—	—	6.97	−.84	2.78	.05
GG....	− .41	− .74	−1.99	.09	−5.35	.94	−3.71	− .10

frequencies of the various judgment sequences which were observed in the experiment reported here. It gives a measure of approximation of the observed value to the value expected by chance. Table I gives the percent of expected values actually observed for Series A, and Tables II and III give the same values for Series D and E respectively. These three tables are identical in form. The first column gives the various possible judgment sequences, the letters L standing for *Lighter*, E for *Equal* and G for *Heavier*. The next four columns are each subdivided in two sub-columns. Each of the four major columns gives the data for a single subject. The first sub-column gives the data for judgment sequences without an interposed judgment; the second sub-column gives the values observed when one undetermined judgment is interposed.

Tables IV, V, and VI are identical in form with Tables I, II, and III. These latter tables deal with the reliability of the difference between observed and expected values in the case of all judgment sequences, and give in all cases the ratio Diff./σ freq.

In addition to these tables a graphical representation of the data of Tables I, II, and III with respect to judgment repetitions has been prepared. Chart A is composed of three sections, one dealing with Lighter judgments, one with Equal judgments, and one with Heavier judgments. Along the absissca will be observed two intervals demarcated by the values 0, 8, and 16, these values indicating the amount of time in beats of the metronome elapsing between comparison and standard weights of proximate pairs. Along the ordinate is to be observed a scale of percentages ranging from 0 to 150. Each subject's graph gives in his distinctive line a picture of the trend of his avoidance of repetition for each category of judgment when the interval between the pairs is increased. Where the line crosses the 100% mark we find, of course, a change from avoidance to preference.

Summary of Results

1. When eight beats of the metronome are interposed between pairs of weights, the tendency against repeating

judgments is still apparent. In the case of every example of possible repetition, a value less than 100 in Table II or a negative value of Diff./σ freq. in Table V is to be observed where members of sequences are temporally proximate.

2. When one judgment is interposed in the data treated in Tables II and V, the tendency against repetition still appears in 8 out of 11 possible instances. In one case only does the result approach statistical reliability, although in two other cases, the value of Diff./σ freq. exceeds 1 in value.

3. Tables III and VI show that when 16 beats are interposed between the presentation of pairs of weights the disposition against repetition is present in all cases but one sequence, and in six cases the value is statistically reliable. As a matter of fact, 9 of the 11 values exceed 1 in magnitude. The chances for all but two out of 11 to be negative occur in random sampling less than once in a hundred trials.

4. When a judgment is interposed within the sequence, the disposition against repetition drops out in the case of Lighter, and in the case of Equal, except in the case of Is, where the value approaches statistical reliability. No clear cut effects exist after a judgment of Greater.

5. Individual differences exist in the extent to which this attitude can be carried over time intervals and over the sensory experience resulting from lifting weights, as well as in the magnitude of the tendency to avoid repetition.

6. From Chart A, as well as Tables I, II, and III, it may be seen that no unanimity of direction of change is shown in the data of four subjects when time is interposed. In the case of Lighter two subjects show a decreasing tendency to avoid repetition while two show an increasing tendency to avoid repetition. In the case of Equals two subjects show a decreasing tendency but Is shows a definite increase at the 16 beat interval. In the case of Heavy we find slight changes in all but subject R. Furthermore, at the 16 beat interval the percentages in many cases are quite small. Six are under 70 percent. One is as low as 25 percent. We see no rapid extinction of the tendency then upon the mere passage of time.

Discussion

The results which have just been reviewed confirm the findings of both Arons and Irwin, and Preston, relative to the disposition against the repetition of judgments. The effect continues to show itself most prominently in the Equal category and least prominently in the Lighter category. Interpolating 8 beats of the metronome appears to weaken it somewhat, but after a 16 beat interval the subjects uniformly show a decided tendency to avoid repetition. What the period of time is which must elapse before the tendency against repetition completely disappears, and whether the apparent weakening at the 8 beat interval is real are matters of speculation. The experimental data in the case of all the Subjects seem to indicate that the attitude may be maintained in considerable force for relatively long periods of time.

The sole criterion available at present for the identification of this blockage with refractory period appears to be the gradual decrease of the blockage and its ultimate disappearance with increasing intervals of time between the members of the series. Examination of the present data does not disclose any marked or significant appropriate change in the magnitude of the values as intervals of time are interposed, which is general throughout the data of all subjects and which leads us to expect its extinction within a relatively short time. Furthermore, not only is the effect carried through interv ls of 16 beats of the metronome, but in the case of Is, for example, we find a measurable effect persisting through an interval of forty beats of the metronome, namely, when sixteen beats of the metronome are interposed between judgments and the repetition of the Equal category is studied with one interpolated judgment (See Tables III and VI). It does not appear possible at the present to explain this phenomenon on the grounds which Dodge and Telford assign.

That the interpolating of the process of comparing a pair of weights in the sequences being studied diminishes but does not entirely destroy the effect of the attitude is again demonstrated. That this should be so is some indication that we are dealing here with a central process, and not one which

may be referred to the simple sensory level. However, none of the experiments to date have established the precise level at which the process may be presumed to operate. Whether it be merely a disposition against using a single name, or against the use of category designations of similar meaning, such as Equal and Same, or whether it transcends such a superficial level is a matter for further investigation.[18]

Conclusions

1. Interpolating rest intervals up to 16 beats of the metronome at 92 beats per minute does not cause the contrast effects due to attitude to decrease to any great extent.

2. Explanation of the avoidance of repetition of the category in the making of psychophysical judgments on the grounds of refractory period is not warranted unless it can be shown that such a period can be influential in voluntary behavior for intervals beyond 26 seconds.

3. It is again shown that the contrast effects exist between members of a series not temporally proximate.

(Manuscript received December 1, 1935)

[18] F. W. Irwin has formulated an experiment to investigate this matter and he and the writer are at work upon it now.

AN EXPERIMENTAL ISOLATION OF SOME FACTORS DETERMINING RESPONSE TO RHYTHMIC CUTANEOUS STIMULATION. I. FREQUENCY, PRESSURE, AND TIME [1]

BY ROGER M. BELLOWS

United States Employment Service, Division of Standards and Research

I. Introduction

The problem is concerned with the isolation of several variables involved in the phenomenon of vanishing tactual flicker. Four variables are controllable and were varied systematically. These are: (1) frequency of intermittent stimulation; (2) pressure, or intensity of the stimulus; (3) time, *i.e.* the length of time the observers had been stimulated by successive stimuli; (4) the temperature of the stimulus; and (5) skin adaptation to temperature. Room temperature and moisture of skin were controlled. A number of other variables could not be absolutely controlled and measured (*e.g.* the *Einstellung* of the observers and their phenomenal criteria of vanishing flicker). These were maintained as constant as possible. We are here concerned with three of these five objectively controllable factors, *viz.* frequency, pressure, and time (adaptation).

The results of the present investigation are related to several extant theoretical issues concerning response to cutaneous stimulation. The following controversial problems are either directly or indirectly implicated in the findings of this study: (1) the layer theory of Head and Rivers, protopathic and epicritic sensibility; (2) the qualitative doctrine of specific nerve energies of Johannes Müller versus the quantitative pattern theory of J. P. Nafe; (3) the relation of cutaneous

[1] This study was conducted under the direction of Professor Robert D. Williams of The Ohio State University, without whose guidance and valuable criticism the work could not have been consummated. The writer is also indebted to the students who gave generously of their time, acting as subjects, during the preliminary and main experiments.

vanishing flicker frequency to discharges in sensory nerve fibers; (4) the learned character of response to tactual stimulation; (5) vibratory versus tactual sensitivity in rapid successive stimulation in teaching the deaf (Gault); and (6) adaptation phenomena.

A. Related Studies

Visual flicker phenomena are not unusual. The spokes of a rapidly turning wheel are observed as a homogeneous surface. The blades of a fan in rapid motion are insensible as separate blades. Moving pictures are commonplace examples of visual flicker phenomena. Naive subjects are readily able to react to visual flicker, Miller (**20**), having recently shown that children react with a fair degree of consistency to a critical frequency of fusion.[2] While there has been much research reported in the literature concerned with visual illusions and a considerable amount dealing with visual flicker, a survey of this literature is considered not particularly pertinent to a study dealing with cutaneous flicker.

Illusions in the cutaneous modality have not played a very important part in psychology. Studies on cutaneous sensitivity appear to make up a rather small proportion of the total volume of articles indexed in *Psychological Index* since 1894.

For the purpose of this work a brief survey is presented of some of the reports of investigations with human subjects. This survey is limited to reports dealing with intermittent tactual and thermal stimulation.

Résumé of Studies on Tactual Flicker.—Halliburton (**15**) makes obscure reference to data on vanishing flicker frequency. He maintains that ". . . vibrations of strings are recognizable as such by the finger, even at a frequency of 1500 vibrations per second. A revolving wheel with toothed edge gives a sensation of smoothness when the teeth meet the skin at the rate of from 480 to 640 per second." No other

[2] In the following pages the phrases—critical frequency, critical frequency of fusion, flicker limen, vanishing flicker frequency, and critical speed of intermittent stimuli—are all used practically synonomously. They refer to that lowest frequency of stimulation at which the observers report that they just perceive continuous pressure.

data or citations are given. Halliburton's obscure reference to flicker phenomena anticipates the basic working hypothesis of Gault.

Gault (**13**) and his students, particularly Goodfellow (**14**), have developed a technique for teaching the deaf by means of tactual stimuli. They maintain that deaf subjects are able to discriminate vibration frequencies as high as 8192 d.v. per second. They found that amplifications must be increased as frequencies are increased; also that the left index finger is superior, in sensitivity to vibro-tactual stimuli, to six other areas of the body investigated.

Karstorf (**19**) investigated cutaneous flicker by a method of intermittent thermal stimuli. He was interested in the critical duration (vanishing flicker frequency) of successive stimulation by means of rays of light which were focussed upon warm spots on the volar surface of the forearm. Karstorf's technique was in one way similar to that of the present investigation in that the durations of his radiated stimuli were equal to the durations of the pauses between stimuli. He found that when the thermal stimuli were presented at a frequency of 2.2 stimulations per second, in which each stimulus lasted .23 second, or at slower frequencies, the stimuli were reported as discontinuous. At faster frequencies his subjects reported continuous warmth sensation. Critical frequency for flicker was discovered to depend upon intensity of the radiated stimuli—the higher the intensity, the higher the vanishing flicker frequency.

Allen and his students in 1924 and 1925 (**2**, **3**) developed a method of studying response to rapid stimulation in which they have been able to control intensity of stimulus, and to measure its duration. The first part of the present investigation (that dealing with the variable of pressure) was performed to check their results, particularly with regard to conclusions concerned with the layer theory. Allen verifies Head's view (**16**, **17**) of protopathic and epicritic sensibility. The apparatus here used is, in essentials, somewhat similar to that used by Allen *et al.* The adaptation (or fatigue) to constant intermittent stimulation which is reported below was not investigated by Allen.

II. Method

A. Apparatus.—The apparatus constructed for investigation of the pressure variable provided intermittent pulsations of air which were directed upon the skin of the subject. It also enabled the experimenter to measure and control the intensity of the stimulus and to record the frequency of the intermittent stimulation. It was possible for either the subject or experimenter conveniently to vary this frequency. For purposes of clarity the description of the apparatus may be divided into three main parts. The first is concerned with the source of the air stimulus and its intensity control, the second with the variation of the speed of intermittent frequency, and the third with the recording of the frequency of vanishing flicker.

The air blast method was adopted for the reasons enumerated by Allen. Allen states, "Obviously no solid point could be used. . . . The advantages of employing a jet of air as a source of interrupted stimuli are: first, its pressure can be definitely controlled and accurately measured . . . ; second it is elastic, and after striking the skin, it rapidly effuses into the surrounding air; third, the pressure with which it strikes the skin, when the latter is at a constant distance from the nozzle, is directly proportional to the pressure in the manometer" (2, 1924, p. 353).

Pressure applied to a solid point could of course be measured, but the stimulus intensity would vary: (1) with distance of the stimulated area from the stimulating point; and (2) with the relative firmness of the cutaneous tissue stimulated. For these reasons apparatus was constructed which would deliver pulsations of air, of known intensity, temperature, and frequency, to the cutaneous surface.

A motor driven air pump or blower of the rotor type provided the source of tactual stimulation. The air was conducted under pressure by a rubber tube to a large tank in the experimental room. This tank acted as a compression chamber which smoothed out the surges in pressure from the rotor pump. The air was then conveyed from the large tank into a small chamber of hard glass located several inches behind the point of stimulation. The temperature of the air (in the experiments in which the air was not heated) was read at this point. From the glass chamber a copper tube passed the air to a mercury manometer. A six volt circuit was so arranged that when closed a relay magnet was energized. This relay opened the A. C. circuit to the motor of the air blower which constituted the air source. Two wire contacts were inserted into the open end of the manometer. One of these contacts was of such length that it always made contact with the top of the mercury column. The other was shorter and fitted with a platinum tip, and could be adjusted to make contact with the rounded top of the mercury in the manometer at any desired height of the column. A millimeter scale was fixed against the side of the manometer. When the upper contact was placed at a certain height corresponding to the desired pressure reading on the millimeter scale the pressure from the air source raised the mercury to this height, at which point the relay circuit was closed by the contact, thus switching off the motor air pump. When pressure diminished in the storage tank the rounded top or meniscus of the mercury column lowered slightly, thereby breaking contact of the relay circuit and starting the motor and increasing pressure to contact. In this way any pressure could be conveniently and accurately (within approximately .01 cm) maintained by adjusting the upper contact to the desired level in the manometer.

The pressures as read in the manometer were calibrated by the following method: the nozzle was detached from its fixed position just back of the sectored disc (there was a clearance of .1 cm) and clamped in position 1.0 cm (the distance from nozzle to skin of subject) above a pair of balance scales; the pressure was then adjusted in the manom-

eter and the air force directed upon the flat surface of one side of the balance. In this way the air blast was measured. The equivalents in grams were obtained for a number of pressures ranging from 0.4 cm to 4.0 cm. In the main experiments reported below all references to intensity are made in terms of centimeters of mercury as indicated by the manometer, which was conveniently placed on the experimental table, so that the pressure could be regulated either by observer or experimenter.

The main feature of the second part of the apparatus is similar to that used by Allen (2). It consisted of a sectored disc which was rotated in front of a nozzle through which the air was forced. The diameter of the aperture of the nozzle was one millimeter. The method of controlling the frequency of the intermittent stimulation here used differed from that used by Allen. His sectored disc "was rotated by a small electric motor (D. C.) whose speed was regulated roughly by a rheostat, and more accurately by a leather brake, in order to adjust it to a point where the intermittent sensations became continuous" (2, 1924, p. 353). The sectored disc was, in the present experiment, turned by means of an almost silent synchronous A. C. motor (1800 R. P. M.). The speed of rotation of the sectored disc was controlled by means of a variable speed reducer of the friction drive type (Central Scientific Company). The gear ratio of this device was controlled by a thumb-screw by means of which the speed of rotation of the sectored disc was adjusted. The disc was divided into twelve open and twelve closed sectors, equally spaced so that one revolution of the wheel gave twelve stimulations. The disc was mounted between the nozzle and the skin of the subject so that a fine stream of air of known pressure must blow through the holes in the disc to impinge upon the subject's skin. The subject, by adjusting the speed of the disc, was enabled to produce any frequency of stimulation varying from 0 to above 500 per second.

The third part of the apparatus consisted of a device for accurately recording the number of revolutions of the disc at any speed chosen by the subject as the point of the vanishing flicker frequency. The recording technique used here differed somewhat from that used by Allen. A contactor was so placed against an eccentric sleeve on the hub of the disc that a circuit was made once for each revolution of the disc. In this circuit was a polygraph magnet which was energized by a six volt storage battery when the contact was made. The armature of the magnet held an ink pen upon the moving tape of the polygraph in such a position that a continuous line was made on the tape. This paper tape of the polygraph moved at a known rate of 4.77 centimeters per second. Thus one complete excursion or cycle of the pen was made by the energization of the polygraph magnet for each complete revolution of the disc. The ink records of the pen excursions made upon the moving tape were counted for the distance of 4.77 centimeters, the result giving the number of revolutions per second. This reading multiplied by twelve gave the frequency of stimulations per second. The reciprocal of this number multiplied by two gave the reading in terms of the duration ($D = 1/2F$) of each discrete stimulus.

An arbitrary assumption is made at this point. It is assumed that each stimulus starts when the edge of a slot of the disc is in front of the nozzle and ends when the other edge of the disc is in front of the nozzle. Thus it is assumed that the stimulus comes on instantaneously, and is turned off abruptly. The slots are, in all of the experiments reported, equal in dimensions to the closed spaces so that duration of 'on' and 'off' phases are equal.

This assumption is of course perhaps not warranted. An attempt was made to take off a record of the pressures for different frequencies on kymograph paper by directing the air jet, while the disc was turning, against a flat surface to which a lever

was attached, its movements being conveyed by a silk thread to a kymographic recording pointer. Various techniques were used but no record of the rapid intermittent air pressures was obtained. For the slow frequencies (below 5 per second) it appeared that the assumption is justifiable.

The skin under the higher pressures was noticeably deformed, *i.e.* a small depression appeared while the area was stimulated. It was noticed under the low power of a binocular microscope that this deformation lasted for some considerable length of time; with four centimeters pressure it could be noticed 60 seconds after the cessation of stimulation. This fact indicates that when the stimulus is, as far as measurement by instrument is concerned, turned off, there may still be afferent impulses initiated in nerve fibers. In view of this fact it appears that if more of the properties of the stimulus could be known we still would not be able to describe its remaining, persistent influence on the skin and the nerve endings which are located there.

The following data are given in terms of stimulations per second as this makes the vanishing flicker frequency more readily comparable to the data given by the electrophysiologists.

B. Preliminary Experiment.—During the preliminary experimentation the writer was interested in obtaining information concerning:

(1) the most efficient method of recording the data;
(2) individual differences in vanishing flicker frequency when objective conditions were maintained constant for all subjects;
(3) the difference, if any, between the readings concerning the frequency of vanishing flicker—
    (*a*) when the experimenter adjusted the frequency for the subject as compared to
    (*b*) readings (taken under the same objective conditions) for which the subject adjusted the speed of rotation of the disc for himself;
(4) the relative consistency of the readings when the frequency of stimuli was—
    (*a*) decreased to the critical point of vanishing flicker from a higher frequency (*i.e.* from phenomenal continuous pressure to sensations of discrete stimuli) as compared to method
    (*b*) in which the frequency was adjusted starting from a slower speed and increased to the critical fusion frequency;
(5) the relative effect of vasoline, for the purpose of preventing drying of the skin, applied to the area of skin stimulated versus the dry skin condition;
(6) whether (under condition 3*a*, above) there was any difference in consistency of the readings when the subject was both blindfolded and when the ears were plugged so that he could have no cues (except perhaps those of a vibratory nature), from other sense modes, as to frequency;
(7) whether it would be necessary to train the subjects to develop an introspective criterion of vanishing flicker or if they could consistently react to the point of critical fusion of stimuli without many practice trials as in visual flicker; and
(8) which area of the skin would give the most consistent readings.

The more essential features of the preliminary experiment are reported in the following paragraphs, in which the findings concerned with items three to eight inclusive in the above list are briefly summarized.

Some differences were experienced by the observers between the readings taken when the experimenter adjusted the frequency as compared to those obtained when the subject produced the frequency. Data were more consistent under the latter procedure. Allen's subject produced their own vanishing frequency. The rate of rotation

of his disc was "adjusted by means of the rheostat and brake, which were operated by the left hand [of the subject]" (2, p. 354). One of the purposes of this experiment was to verify Allen's results on intensity. Since readings are different under the two methods, on account of the added reaction time of the experimenter when he adjusted the frequency, the subjects in the main experiments reported below controlled frequency.

The direction of change of speed (whether the flicker limen was approached from above or below) was found to be a factor. The subject's verbal reports indicate that when he started with a much higher frequency and reduced speed of rotation down to the flicker limen, he always tended to go beyond the critical point and it was necessary to increase frequency to the point of critical fusion. This method required more time for the subject to find the frequency. Early in the preliminary work it was found that time of stimulation tended to lower the vanishing flicker frequency. More consistent results were therefore expected, and found, when the subject approached the limen from a slower speed at which pulsations were clearly discrete and discontinuous. This latter method was followed in the main experiments. The data indicate that when speed is decreased from phenomenally continuous to slower frequency the appearance of flicker is found at a higher rate than when frequency is increased to critical fusion of successive stimuli.

Allen found a second fusion point for flicker. He says, "On further increasing the speed a second point was reached where there was apparent fusion of the stimuli. This second fusion point is not of the same nature as the first; it was less satisfying, so to speak, but still it was a definitely measurable point of fusion" (2, p. 355). We did not find it definitely measurable.

Drying of the skin and control of humidity were considered. For the purpose of preventing drying of the cutaneous surface by the air blast in the main experiments, a thin covering of petroleum jelly was spread evenly over the area stimulated before each sitting. This procedure tended to mitigate any possible influence of the differences in humidity of the air stimulus. For this reason it was not considered practicable to control the possible variable of humidity.

For the purpose of obtaining information concerning the influence of drying of the skin the following technique was devised. One subject (the writer) stimulated his skin with the designated pressure for a period of two minutes before the first trial, then took a reading, and subsequently took readings at intervals of two minutes during which the area was constantly stimulated.

The decrease in frequency with time was at first attributed to drying influences. Allen reports somewhat longer durations (lower frequencies) with the dry skin as compared to vasoline coated skin. At 1 cm his subject, Hollenberg (stimulating the right index finger), reported critical duration at .010 seconds; at 2 cm, .0074; at 3 cm, .0058. When vasoline was used Hollenberg's durations were decreased as follows: for 1 cm, .0089; at 2 cm, .00628; at 3 cm, .00475; and at 3.8 cm at .0039 (2, p. 355).

But it was thought later that drying influences were not alone sufficient to cause the drop in frequencies with stimulation. The preliminary data furnished the cue for the investigation of adaptation which is reported below. In this investigation a thin film of vasoline was spread over the cutaneous area stimulated.

It was found that subjects reacted relatively consistently when visual and auditory cues were eliminated as when able to see and hear during the experiments. The elimination of auditory cues was made by inserting rubber plugs, such as are used by divers, well into the external meatus of the ear. For the elimination of the visual modality the subjects were blindfolded.

It was found necessary in the case of all subjects to train for a certain degree of

consistency of readings. In visual flicker it does not seem necessary to train subjects. We attempted to attain a point of proficiency at which there was 10 percent or less variation in number of stimuli per second for vanishing flicker frequency under controlled conditions of temperature, pressure, and time. In the case of subject B (the writer) the practice period consumed five weeks of almost daily practice periods of thirty minutes each. All of the seven o's, undergraduate students, practiced for the same length of time for twelve practice periods before their records were sufficiently consistent to justify them in starting the experimental program. Some of the seven [3] subjects who were tried during the training period were either too inconsistent (more then 10 percent variation) to be of value, or they were unable to serve for various reasons, during the principal experiments, which were performed at a later date.

The frequency of vanishing flicker is not a point (notwithstanding the fact that this term is conventionally applied to the concept), but a range, and on this account some variation is to be expected on the ground of observational error. Upon increasing the speed of the wheel the stimulations, which at first are perceived as discrete puffs, gradually fade into a smooth stream, the sensation being one of continuous temperature and continuous light pressure only. The subjects were instructed to respond by pressing a key (which started the polygraph motor and thereby moved the tape, and also made a potential circuit through the polygraph magnet) when the speed of vanishing flicker frequency was reached. Since there is a tickle sensation which sometimes accompanies the sensation of continuous pressure and temperature above the frequency for critical fusion, some confusion was experienced at first. After training this tickle, when present, was disregarded.

Various areas were tried in the preliminary experiments. The main problem here was to find the most consistent area to work on with the problem variables of pressure, temperature, etc. We were not particularly interested in the relative critical frequencies for the different areas, although some data, of a tentative nature only, were obtained on this phase of the problem. Areas at different locations on the arm, the back, neck, finger tips, and lips were tried. Observers responded in general with somewhat higher frequencies to stimulation on the lips, and it was convenient to stimulate here for several reasons. They reported greater introspective certainty and sureness of their criterion on this area.

The subjects were instructed to always present the most sensitive surface to the air stream (touch, temperature spots) by slight movements of the head—the lips are known to be extremely sensitive (short two point limen and numerous temperature spots) so that we feel justified in selecting this surface as the experimental area.

The subjects were seated comfortably before the sectored disc so that the center of the lower lip (in all cases except that of subject C, in which the upper lip was stimulated) was in front of the stimulating air jet. The head was held in a special head rest so constructed that the sides of the support could be adjusted firmly to the sides of the head. The jet of air was directed perpendicularly upon the surface of the lip which was coated with a light film of petroleum jelly to prevent drying and to mitigate the differential influence of the factor of humidity of the air blast, as mentioned above. The lip was held in all cases approximately one centimeter from the opening of the nozzle.

C. Method for Determining the Influence of Pressure and Adaptation upon Vanishing Flicker Frequency.—A stop watch was placed on a ring stand in front of the subject's eyes for the work on adaptation to prolonged intermittent stimulation. The motor

[3] Two of these seven o's participated with three others in an experiment in which the temperature variable was isolated—these data are not here reported; two were eliminated; three served in the part of the investigation reported here.

air pump was started and the pressure column adjusted (and automatically maintained) to the desired height in the mercury manometer. The subject was instructed to maintain the speed of the sectored disc at the frequency of critical fusion for twenty minutes at each sitting. The frequency of stimulation was adjusted by means of the thumbscrew operated by the right hand of the subject. The first reading was taken by touching the recording key with the fingers of the left hand at the end of the first minute. At the end of each successive minute for 20 readings during each sitting the subject pressed the key with his left hand, thereby taking one reading. Each subject sat five different times for each of the five pressures—1.0 cm, 1.5 cm, 2.0 cm, 2.5 cm, and 3.0 cm. The data for the pressure variable were taken from the initial readings at each of these five pressures. The rest of the data which are shown below in Figure 6, constitute the results for the time variable.

III. Results

The data obtained in the main experiment are presented in the figures. These investigations are divided into the following two parts: (*A*) the relation of pressure and vanishing flicker frequency, in which the stimulus temperature was (nearly) the same as the room temperature; (*B*) the influence of time of intermittent stimulation upon vanishing flicker frequency, when stimulus temperature was (approximately) equal to room temperature, and under different conditions of pressure which was varied from 1.0 cm to 3.0 cm inclusive in units of .5 cm.

A. The Pressure Variable

Figures 1 and 2 show the data obtained in the experiment upon stimulus intensity. Five trials were taken for each of three subjects and for five stimulus intensities. These trials consist of the readings taken at the end of the first minute of the time of intermittent stimulation experiment. Stimulated by a pressure of successive stimuli of 1.0 cm subject B reported frequencies of vanishing flicker of 51.6, 50, 72, 60, and 66 for the five different sittings—the mean is 59.9 with a mean deviation of 7.3. The frequencies taken for the same subject for 1.5 cm, 2.0 cm, 2.5 cm, and 3.0 cm, and for the other two subjects under like experimental conditions, are shown in Fig. 1, in which frequency of stimulation is indicated on the ordinate while pressures are shown on the abscissa. The heights of the bars (Fig. 1, *G*) which intersect the curves are equal to the magnitude of one mean deviation of the readings represented at that point.

In Fig. 1 the curves for subjects C and B (broken lines) are superimposed upon the curve for subject G (the bold solid line) in order to bring out the individual differences. It is

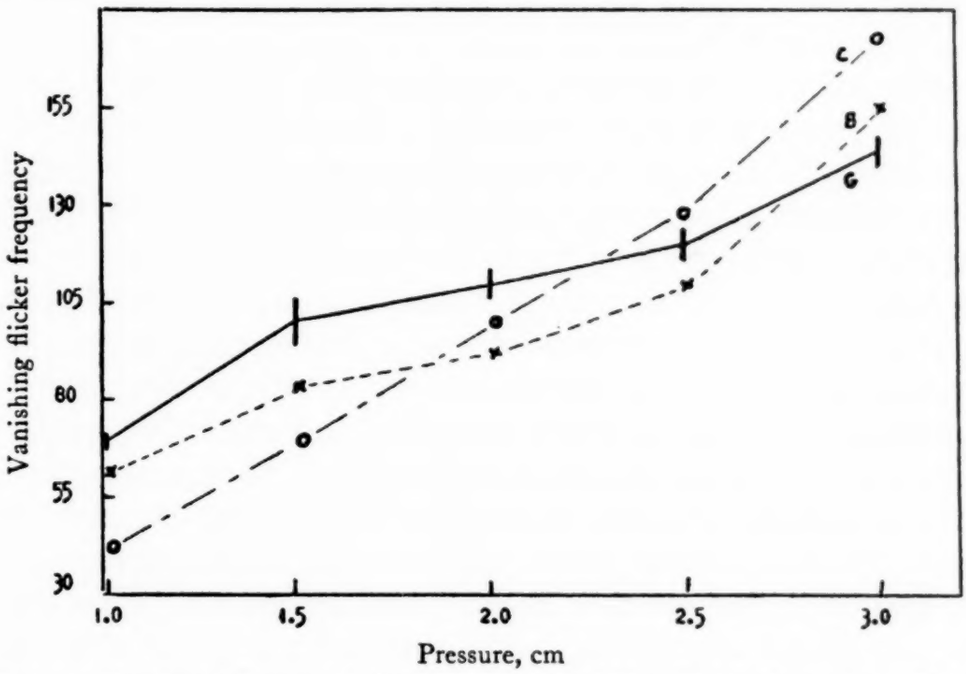


FIG. 1. Showing influence of pressure upon vanishing flicker frequency. Height of the bars indicate magnitude of mean deviations.

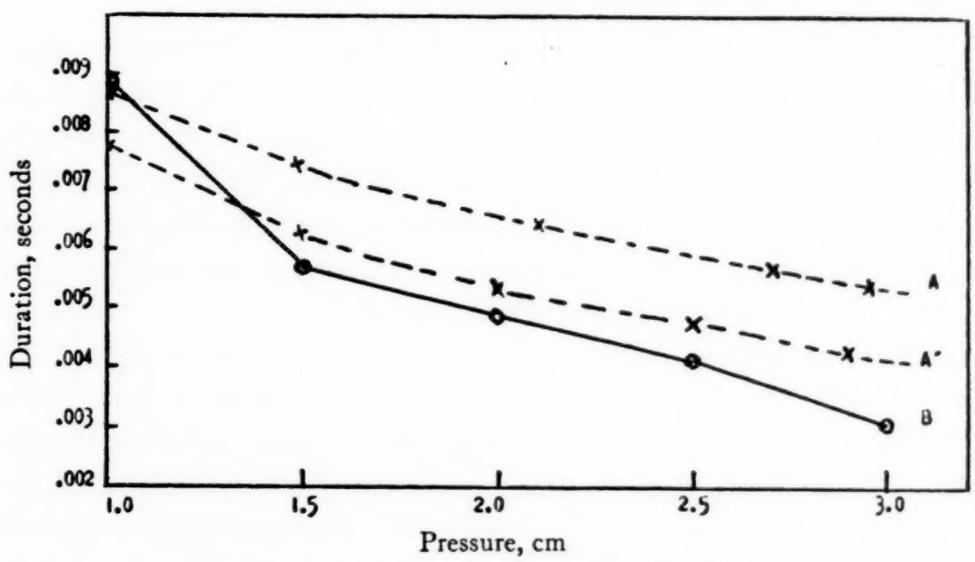


FIG. 2. Pressure and duration, all subjects.

apparent that vanishing flicker frequency increases, rather uniformly in every case (when the averages of five readings are used), with an increase in pressure between 1.0 cm to

3.0 cm. The influence of pressure change seems greater for subjects B and G between the intervals 1.0 cm to 1.5 cm than for 1.5 to 2.5 cm. Similarly, and for all subjects, the frequency change from 2.5 to 3.0 cm is greater than that for the middle range of intensities used.

Figure 2 shows the same data in terms of the critical duration of each individual stimulus. The full line shows the readings taken in the present experiment while the broken lines indicate the readings tabulated by Allen and Weinberg (**3**); (they obtained two critical frequencies for each pressure). In both sets of data it is noted that the same general trend is present—duration decreases with increase in pressure.

Curve *A* in Fig. 2 represents the curve of Allen for his superficial sensitivity while curve *A'* indicates the relation which he found for his deep sensitivity. Allen states ". . . both sensations are seen to behave in the same manner, the deep, however, being more sensitive than the superficial." He identifies two series of fusion points, the first "represents the superficial system of tactile sensibility, . . . the second series of fusion points, represents the deep system" (**3**, p. 379). In the present study only one series of points was found; these are represented in Fig. 2 by curve *B*.

Large individual differences of influence of pressure upon frequency are evident. For example at 1.0 cm pressure the average frequency of subject G's five trials is 67.5 as compared to 40.8 for subject C and 60.0 for subject B. G's curve starts out at a higher level but flattens out toward the center, remaining lower than that of the others for the higher pressures. These individual differences indicate that either (1) the phenomenal criterion of vanishing flicker shifts differently for the different subjects, or (2) that differences in sensitivity (receptor, neurological differences) exist from subject to subject. Since the discrimination of cutaneous vanishing flicker is at least to some extent a learned phenomenon we could perhaps expect such differences on one basis or the other, or both.[4]

There is no general trend of difference of scatter for the

[4] It will be noted in a later paper that temperature variation of the stimulus has a greater influence upon the readings of subject G than upon other subjects.

different subjects. In Fig. 1 the measure of dispersion is shown (for subject G only) by the vertical bars. The data for subject G show a mean deviation of 2.0 for the pressure of 1.0 cm while subject C has an M.D. of 4.8 and B of 7.3. For the next pressure, 1.5 cm, this trend is reversed, G's M.D. is 11.6, that for C is 6.3, and for B, 2.0. The average of the 5 mean deviations for the different subjects for all pressures are: subject B, 7.6; subject C, 10.3; and subject G, 6.53. There is a clear tendency in the case of subject C for the mean deviations to increase as pressure increases. This tendency appears to exist, but is not as obviously uniform, in the case of other subjects.

Since we would expect the mean deviations to increase somewhat as the means increase with the higher pressures the mean deviations have been reduced to percents for the purpose of determining any tendency for comparatively greater or less variability for the higher pressures. For subject B these percents, for the pressures of 1.0, 1.5, 2.0, 2.5, and 3.0 cm, are: 12, 2, 15, 7.5, and 8 respectively. For subject C they are 10, 12, 9.4, 12, and 9; and for subject G, 3, 11, 4, 6, and 4 respectively. It is apparent that there is some tendency for variability of readings to decrease relatively slightly at the higher pressures but this tendency is neither marked nor uniform.

B. The Time Variable

The data taken in this part of the experiment are presented diagrammatically in Figs. 3, 4, 5, and 6. The subjects

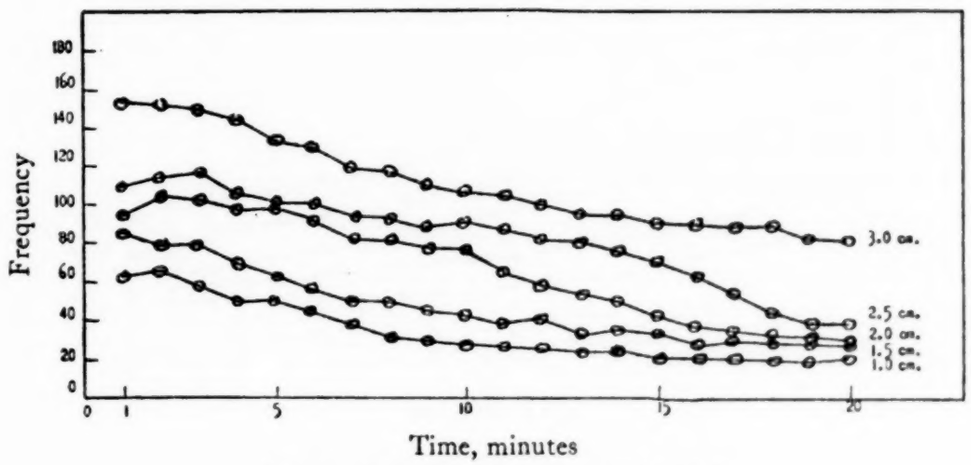


FIG. 3. Time and frequency, subject *B*.

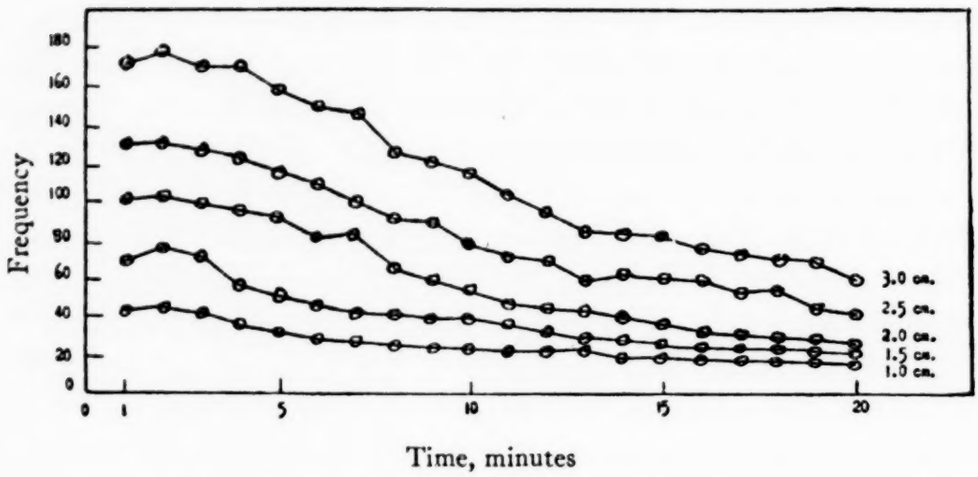


FIG. 4. Time and frequency, subject *C*.

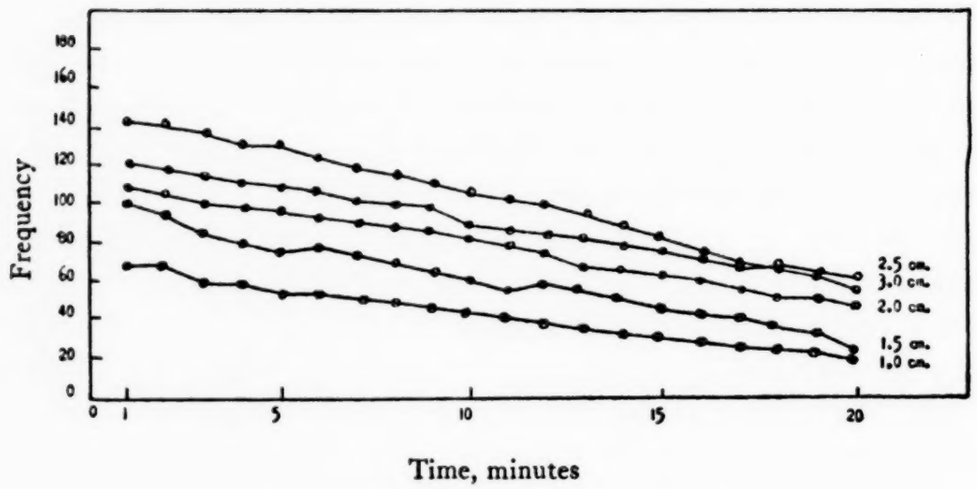


FIG. 5. Time and frequency, subject *G*.

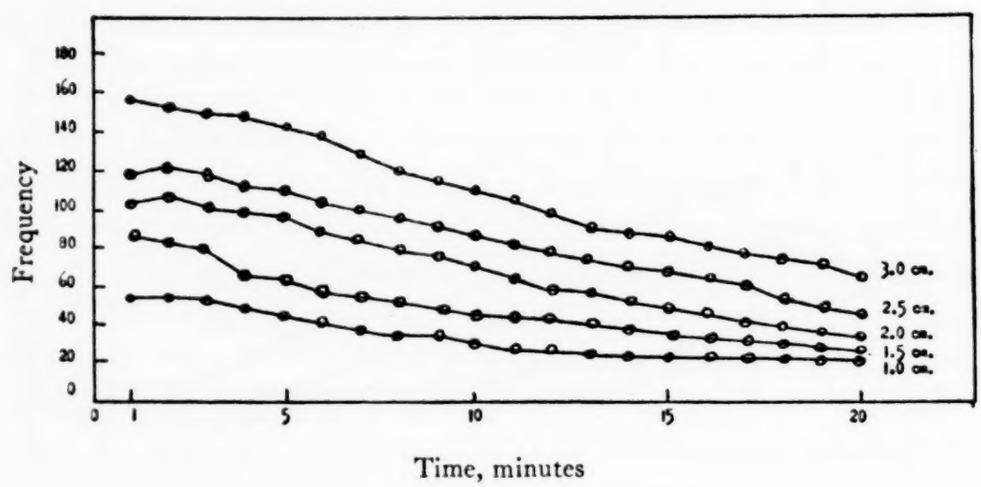


FIG. 6. Time and frequency, all subjects.

were seated before the instrument and instructed to maintain the frequency at the speed of vanishing flicker for a period of 20 minutes, and to touch the recording key at the end of each consecutive minute.

Figure 3 shows these data for subject B. Vanishing flicker frequency is indicated on the ordinate while time in minutes is shown on the abscissa. The upper curve in the figure shows the drop in frequency with time of intermittent stimulation for the pressure 3.0 cm. The next curve indicates the trend for the pressure of 2.5 cm, etc. Figures 4 and 5 show the same general tendencies for the subjects C and G, respectively. Figure 6 schematically recapitulates the entire data taken in this part of the experiment—the average of all of the readings for all subjects under the different conditions of pressure and time.

IV. Discussion

Three main tendencies are noted by inspection of the data presented in the tables and figures.

First, it appears that frequency decreases with time (adaptation) in a fairly regular and uniform manner. The function is almost linear. There is a slight tendency for the effect of time to be mitigated toward the thirteenth, fourteenth, and fifteenth minute. The curves in general appear to flatten out near these points although this tendency is not universal (*e.g.* subject B, 2.5 cm, and the curves of subject G).

Secondly, there is a tendency for some of the readings to increase at the second (or third) minute and then decrease. This enhancement is evident in two of the curves for subject B (Fig. 3), in all of those for subject C (Fig. 4), and in three of the curves in Fig. 6.

Thirdly, frequency decreases with time more, proportionately, for the lower pressures than for the higher ones. The data shown for all subjects indicate that for the pressure of 3.0 cm the initial frequency was decreased by 25 percent of the initial frequency by the end of the eighth minute.

Protocols of the subjects verbal reports taken during the course of this experiment indicate that from the eighth to thirteenth minutes subjects would feel 'numbness' and periods

of 'less sensitive' feeling. Inspection of the standard deviations (not here shown), however, indicates that readings do not scatter proportionately more during this period than for other periods. On the contrary, scatter of readings seems to be greater for the first five minutes than during any other period.

V. Conclusions

Several conclusions seem warranted. These are:

1. Certain relationships between the variables of frequency of skin stimuli, their intensity, and time of stimulation, are fixed. It has been found that (*ceteris paribus*):

(*a*) Intensity varies directly with frequency,
(*b*) Time varies directly with intensity, and
(*c*) Frequency varies inversely with time.

2. Only one frequency of vanishing flicker was found under constant objective conditions with regard to the variables of pressure, adaptation time, and temperature. No specific classification of epicritic-protopathic or superficial-deep sensitivity appears to be necessary for interpreting the results.

3. The data are perhaps more readily interpreted in terms of the quantitative pattern theory of cutaneous phenomena than on the basis of the qualitative, peripheral specificity doctrine.

(Manuscript received January 11, 1936)

VI. Bibliography

1. Arps, G. F., Introspective analysis of certain tactual phenomena, *Psychol. Rev.*, 1912, **9**, 337–351.
2. Allen, F., and Hollenberg, A., On the tactile sensory reflex, *Quart. J. Exper. Physiol.*, 1924, **14**, 352–378.
3. Allen, F., and Weinberg, M., The tactile sensory reflex, *Quart. J. Exper. Physiol.*, 1925, **15**, 377–383.
4. Allen, F., and Macdonald, P. A., The sensations of temperature, pain, and pressure, *Quart. J. Exper. Physiol.*, 1927, **16**, 321–332.
5. Bazett, H. C., and McGlone, B., Note on pain sensations which accompany deep puncture, *Brain*, 1928, **51**, 18–23.
6. Bartlett, F. C., Adaptation et fatigue, *J. de. psychol.*, 1933, **30**, 673–691.
7. Burtt, H. E., Tactual illusions of movement, *J. Exper. Psychol.*, 1917, **2**, 371–385.
8. Burtt, H. E., Cutaneous space, *Psychol., Bull.* 1922, **19**, 338–342.
9. Cattell, M., and Hoagland, H., Response of tactile receptors to intermittent stimulation, *J. Physiol.*, 1931, **72**, 392–404.

10. DRURY, M. B., AND DALLENBACH, K. M., The response of cold spots under successive stimulation, *Amer. J. Psychol.*, 1931, **43**, 490–496.
11. EARHART, R. H., AND DALLENBACH, K. M., The response of warm spots under successive stimulation, *Amer. J. Psychol.*, 1933, **45**, 722–729.
12. FREDERICQUE, H., Testing excitability by means of a time factor, *Physiol. Rev.*, 1928, **8**, 541–544.
13. GAULT, R. H., The American Institute for the deaf-blind, *Rehabilitation Rev.* 1933, **12**, pp. 8.
14. GOODFELLOW, L. D., The sensitivity of the finger-tip to vibrations at various frequency levels, *J. Franklin Institute*, 1933, **216**, 387–392.
15. HALLIBURTON, W. D. *Handbook of physiology.* Philadelphia: Blakiston's, 1917, 12th ed., pp. xiv + 930.
16. HEAD, H., AND SHERREN, J., The consequences of injury to the peripheral nerves in man, *Brain*, 1905, **28**, 116 ff.
17. HEAD, H., RIVERS, W. H. R., AND SHERREN, J., The afferent nervous system from a new aspect, *Brain*, 1905, **28**, 99–115.
18. HOAGLAND, H. *Pacemakers in relation to aspects of behavior.* New York: Macmillan, 1935, pp. x + 138.
19. KARSTORF, F., Ueber die Verschmelzung der Warmempfindung bei rhythmische erflogenden Reizen, *Zeitsch. Biol.*, 1920, **71**, 1–18.
20. MILLER, V. L., *Studies on the critical frequency limen for visual flicker in children.* Ph.D. Dissertation, The Ohio State University, 1931.
21. NAFE, J. P. The pressure, pain and temperature senses. In *A Handbook of General Experimental Psychology.* Worcester, Mass.: Clark University Press, 1934, pp. 1037–1087.
22. RENSHAW, S., WHERRY, R. J., AND NEWLIN, J. C., Cutaneous localization in congenitally blind versus seeing children and adults, *J. Genet. Psychol.*, 1930, **38**, 239–248.
23. THALMAN, W. A., The after effect of movement in the sense of touch, *Amer. J. Psychol.*, 1922, **33**, 268–276.
24. WEBER, E. H. Tastsinn und Gemeingefühl. In Rudolph Ostwald's *Klassiker der Exakten Wissenschaften*, Nr. 149, Leipzig: 1905. Offprinted from Wagner's *Handwörterbuch*, 1846, 480–588.

RELIABILITY AND VALIDITY OF EYE-MOVEMENT MEASURES OF READING *

BY MILES A. TINKER

University of Minnesota

Since the turn of the century eye-movement records have been employed more and more as measures of reading performance. There are now several important laboratories equipped for studying eye movements by the photographic technique. In an earlier paper, Tinker [1] suggested that the wide-spread use of eye-movement measures of reading ability justifies an evaluation of the photographic technique. It is time to inquire whether photographic records are reliable and valid measures of reading.

Until quite recently investigators have apparently assumed that the various measures of eye movements were sufficiently reliable to justify rather definite conclusions from the data collected. Recently published reliability coefficients are high. Litterer [2] obtained coefficients of .91 and .92 for perception time and fixation frequency on easy prose, and .91 and .78 for the same measures on scientific prose. Also using adult subjects, Eurich's [3] reliability coefficients were: fixation frequency = .85; pause duration = .91; regression frequency = .77. With children [4] the three comparable reliability coefficients for reading three short paragraphs were .91, .86, and .91. For any single paragraph the coefficients ranged

* The expenses of this study were met by a research grant from the Graduate School, University of Minnesota.

[1] Tinker, M. A., Use and limitations of eye-movement measures of reading. *Psychol. Rev.*, 1933, **40**, 381–387.

[2] Litterer, O. F., An experimental analysis of reading performance. *J. Exper. Educ.*, 1932, **1**, 28–33.

[3] Eurich, A. C., The reliability and validity of photographic eye-movement records. *J. Educ. Psychol.*, 1933, **24**, 118–122.

[4] Eurich, A. C., Additional data on the reliability and validity of photographic eye-movement records. *J. Educ. Psychol.*, 1933, **24**, 380–384.

from .62 to .87. Similarly, Frandsen [5] obtained reliability coefficients ranging from .78 to .88 for adult reading of prose, and from .55 to .90 for reading objective questions. All the above coefficients were obtained by the odd-even or split-half method. Test-retest reliabilities are needed.

There are several aspects to be considered in studying the validity of eye-movement measures of reading. These have been summed up by Tinker.[6] An evaluation of the measures revealed that perception time and fixation frequency are good measures, regression frequency a fair, and pause duration a poor measure of reading speed. It was further shown that the direct method of observing eye movements is quite inaccurate and only useful to reveal general trends.

Validity of eye-movement measures has been computed by correlating the measures with achievement on standardized reading tests. By this method Litterer [7] obtained for adults coefficients ranging from $-.27$ to $-.51$. (The negative coefficients indicate that the reading score increased as the fixation frequency decreased, etc.) Comparable results from Eurich [8] ranged from $-.02$ to $-.24$ for adults, and from $-.11$ to $-.55$ for children. It is clear from these coefficients that validity with achievement test scores as criteria is low. As will be shown later, this is an unsatisfactory technique of determining validity. Other methods of attacks should be employed.

The purpose of the present study is twofold: (1) To determine by the test-retest method the reliability of eye movements in reading materials of various lengths. (2) To investigate the validity of eye-movement measures by studying adaptation to the artificial laboratory situation, and by computing validity coefficients (*a*) with scores on standardized reading tests as criteria and (*b*) with performance on strictly comparable reading material as criterion.

[5] Frandsen, A., An eye-movement study of objective examination questions. *Genet. Psychol. Monog.*, 1934, **16**, No. 2, 79–138.

[6] Tinker, M. A., Eye movement, perception and legibility in reading. *Psychol. Bull.*, 1936, **33**, 275–290.

[7] Litterer, O. F., *op. cit.*

[8] Eurich, A. C., *op. cit.*, p. 118 and p. 380.

A photographic technique[9] was employed to record eye movements during the reading of three kinds of material. (1) Seven paragraphs of connected discourse normally adapted to about fourth-grade reading. This yielded a measure of speed with comprehension practically constant at an easy level. The number of lines in the succeeding paragraphs was: paragraph P, 8 lines; A, 5; B, 10; C, 8; D, 12; E, 6; F, 5. (2) Seven successive paragraphs of relatively hard scientific prose taken from Holmes' *The evolution of animal intelligence.* The number of lines was: in paragraph A, 7 lines; B, 6; C, 11; D, 15; E, 16; F, 6; G, 16. In addition there were two introductory paragraphs of this material to orient the reader in the subject matter before the eye movements in reading the succeeding sections were photographed. (3) Two sections (14 lines each) of four short paragraphs taken from the Chapman-Cook Speed of Reading Test. These were taken from a part of the test not previously read by the subject.

The easy and difficult prose selections were printed on egg-shell paper stock, in 10 point Scotch Roman type with 2 point leading and a 25 pica line length; the speed of reading test, in 10 point Antique set solid with a 19 pica line length. In addition to the reading done before the camera, scores were obtained for each subject on the following tests: Test I (vocabulary) of the Minnesota Reading Examination for College Students, Form A; Test I (paragraph meaning) of the Iowa Silent Reading Tests, Advanced Test, Form A; Chapman-Cook Speed of Reading Test, Form A.

The following procedure was adopted. The performance tests were given to the subjects in regular class groups. The subjects then went to the photographic laboratory for two sittings separated by one or two days. At the first sitting the subject was told how the photographing was done. Then photographs were taken while the subject read the first four paragraphs of the easy material. Thus paragraph P (practice) gave a record of eye movements on first exposure to the apparatus. The two preliminary paragraphs of the difficult prose were then read for orientation to the material. This was followed by photographs for sections A, B, C, and D. Finally photographs were taken while reading the first Chapman-Cook selection (CC1). On the second day the subjects were oriented by rereading the last paragraph done at the preceding sitting. This was followed by photographing the remaining three sections of easy material, the three of hard prose and the second Chapman-Cook selection. Comprehension questions were asked on all selections read. Seventy-seven university sophomores who were taking elementary psychology served as subjects in the experiment.

To further study the degree of adaptation to the experimental situation, 57 new subjects (freshmen) were brought to the laboratory individually. Each was given Form A of the Chapman-Cook Speed of Reading Test at a table, which is the customary way of giving the test. Then, after explaining how the apparatus works, the subject was seated at the camera, the light and focusing adjustments made, and the motor started. After a short practice trial, Form B of the Chapman-Cook Test was read. The subject thought he was being photographed but in reality there was no film in the camera. Scoring was done as on Form A. A time limt of 1¾ minutes was used. Comparison of scores on Forms A and B should reveal the adaptation to the experimental situation.

[9] Tinker, M. A., Apparatus for recording eye movements. *Amer. J. Psychol.*, 1931, **43**, 115–117.

The Scores

The mean scores, the S.E.'s of the means, and the reliabilities[10] of the achievement tests are listed below:

C.C. Speed R.: $M = 21.8$, $\sigma_M = .43$, Reliability $r = .76$
Iowa Silent R.: $M = 59.7$, $\sigma_M = 1.51$, Reliability $r = .87$
Minn. R.E. Vocab.: $M = 53.1$, $\sigma_M = 1.48$, Reliability $r = .90$

The averages and distributions are typical and the reliabilities satisfactory.

In Tables 1, 2, and 3 are given the means and S.E.'s of the means for the eye-movement measures for each type of material read before the camera. The means for perception time, fixation frequency, and regression frequency are averages per line; for pause duration the average is for all pauses in the selection. The unit of measurement for pause duration and perception time (number of fixations times pause duration) is one-fiftieth second.

Examination of the tables reveals some variation from one paragraph to another within each kind of material. This is especially noticeable in paragraph A of the hard prose in comparison with the succeeding sections. Apparently it is more difficult to adapt to the requirements of reading hard prose for comprehension than with easy prose. Also it is possible that

TABLE 1

MEANS AND S.E.'S OF THE MEANS FOR EYE-MOVEMENT MEASURES FOR EASY PROSE

Paragraph	No. Lines	Perc. Time		Fixations		Regressions		P. Duration	
		M	σ_M	*M*	σ_M	*M*	σ_M	*M*	σ_M
P	8	90.78	3.26	8.51	.27	1.19	.10	10.79	.19
A	5	97.19	2.84	8.83	.20	1.35	.09	11.07	.18
B	10	80.36	2.21	7.52	.15	1.06	.08	10.75	.16
C	8	77.04	2.10	7.22	.18	0.94	.07	10.85	.16
A–C	23	82.65	2.22	7.65	.16	1.04	.07	10.87	.15
D	12	90.00	2.09	8.12	.17	1.19	.08	11.05	.16
E	6	79.48	2.17	7.50	.19	0.92	.08	10.53	.14
F	5	90.65	2.26	8.73	.19	1.39	.11	10.85	.17
D–F	23	88.57	2.05	8.14	.17	1.19	.08	10.91	.15
A–F	46	85.68	1.97	7.89	.16	1.13	.07	10.89	.16

[10] Litterer, O. F., *op. cit.;* and Litterer, O. F., An experimental study of visual apprehension. *J. Appl. Psychol.*, 1933, **17**, 266–276.

TABLE 2

MEANS AND S.E.'S OF THE MEANS FOR EYE-MOVEMENT MEASURES FOR HARD SCIENTIFIC PROSE

Paragraph	No. Lines	Perc. Time		Fixations		Regressions		P. Duration	
		M	σ_M	*M*	σ_M	*M*	σ_M	*M*	σ_M
A	7	130.32	3.54	10.73	.24	2.22	.11	12.16	.20
B	6	102.27	2.87	9.20	.21	1.42	.09	11.24	.19
C	11	94.48	2.32	8.45	.18	1.31	.09	11.25	.09
D	15	107.21	2.56	9.25	.19	1.66	.09	11.59	.18
A–D	39	106.82	2.51	9.27	.18	1.74	.08	11.54	.17
E	16	99.09	2.15	8.83	.18	1.55	.09	11.23	.14
F	6	103.31	2.32	9.26	.18	1.72	.09	11.11	.18
G	16	109.57	2.30	9.41	.19	1.72	.09	11.96	.16
E–G	38	104.27	2.08	9.15	.17	1.66	.08	11.32	.15
A–G	77	100.10	2.08	9.10	.16	1.61	.08	11.54	.16

TABLE 3

MEANS AND S.E.'S OF THE MEANS FOR EYE-MOVEMENT MEASURES FOR CHAPMAN-COOK SPEED OF READING TEST

Section	No. Lines	Perc. Time		Fixations		P. Duration	
		M	σ_M	*M*	σ_M	*M*	σ_M
CC_1	14	64.84	2.26	5.89	.18	10.99	.20
CC_2	14	69.39	2.54	6.45	.22	10.78	.17
1–2	28	67.17	2.20	6.17	.19	10.89	.18

paragraph A is more difficult than the following ones. Pause duration in the speed test is quite comparable to that in the easy prose. The other measures are not comparable since the line length is shorter in the speed test. As one would expect the easy materials were read in much shorter times than the more difficult prose.

RELIABILITY

Previously cited reliability coefficients for eye-movement measures have been high. They have been computed, however, by the odd-even or split-half method. They signify, therefore, only the consistency of performance during a single sitting. In the present experiment reliability was investigated by correlating performance on one paragraph with performance on subsequent paragraphs (1) read at the same sitting

and (2) read at a second sitting. Analysis of the reliability coefficients will give information on the following points: (1) the effect of habituation to the reading situation and reading materials upon reliability; (2) the consistency of performance from one paragraph to another at the same sitting, and consistency from one sitting to the other; (3) the effect of length of reading copy upon reliability, and the comparative reliability of the various measures.

The reliability coefficients are given in Tables 4, 5, 6, and 7. In addition the coefficients for the speed of reading test, selection CC1 versus CC2, are .72 for perception time, .65 for fixation frequency, and .81 for pause duration.

In Table 4 we find that perception time for the first paragraph read, the practice (P) selection, is nearly as reliable as subsequent paragraphs. Evidently most readers, after the experimenter has given a description of how the apparatus

TABLE 4

RELIABILITY OF PERCEPTION TIME AS A MEASURE OF READING PERFORMANCE

	Easy Prose				$N = 77$				
	First Sitting				Second Sitting				A 11
Selection	A	B	C	A–C	D	E	F	D–F	A–F
P	.76	.71	.70	.74	.65	.51	.68	.66	.73
A		.86	.77	.91	.69	.77	.83	.78	.90
B			.79	.95	.69	.71	.78	.79	.88
C				.92	.65	.74	.78	.77	.85
A–C					.70	.79	.83	.82	.95
D						.70	.82	.93	.85
E							.79	.88	.88
F								.91	.91
D–F									.96
	Hard Prose				$N = 77$				
	B	C	D	A–D	E	F	G	E–G	A–G
A	.73	.74	.69	.87	.51	.55	.52	.56	.77
B		.82	.71	.87	.68	.64	.61	.70	.81
C			.76	.91	.63	.67	.61	.68	.84
D				.88	.52	.54	.60	.63	.81
A–D					.65	.65	.66	.72	.90
E						.77	.73	.91	.81
F							.74	.87	.78
G								.93	.88
E–G									.87

works, settle quickly into consistent and typical reading. Observation revealed certain subjects, however, who were disturbed by the bright light shining into their eyes and by the rigid confinement in the headrest. They need practice as well as explanations to bring adequate adaptation to the artificial reading situation at the camera. It is probably best to give all subjects a practice trial to promote adaptation to the set-up.

TABLE 5

RELIABILITY OF FIXATION FREQUENCY AS A MEASURE OF READING PERFORMANCE

	Easy Prose				$N = 77$				
	First Sitting				Second Sitting				A 11
Selection	A	B	C	A–C	D	E	F	D–F	A–F
P	.78	.79	.73	.81	.76	.59	.79	.79	.82
A		.84	.72	.89	.70	.73	.75	.78	.86
B			.80	.98	.80	.75	.75	.85	.90
C				.90	.74	.76	.76	.81	.89
A–C					.82	.80	.80	.88	.96
D						.75	.78	.95	.91
E							.73	.89	.88
F								.87	.88
D–F									.97
	Hard Prose				$N = 77$				
	B	C	D	A–D	E	F	G	E–G	A–G
A	.74	.68	.74	.85	.58	.57	.57	.59	.77
B		.75	.80	.87	.70	.68	.65	.73	.82
C			.80	.91	.69	.73	.68	.76	.84
D				.93	.65	.63	.66	.72	.84
A–D					.74	.72	.71	.78	.90
E						.80	.77	.93	.84
F							.75	.87	.80
G								.93	.86
E–G									.90

In general there is a slight tendency for the reliability coefficients to increase with continued practice before the camera. Part of the increase may be due, however, to adaptation to the reading material itself. These trends are discovered by examining the correlations of each selection with the following selection. The variation of the coefficients in progressing downward in columns 5, 9, and 10 also show the

same trend. The spurious factors in these coefficients are fairly constant (.07 to .10 points) and do not obscure the trends due to practice. The coefficients in these three columns show that the paragraph preceded by most practice

TABLE 6

RELIABILITY OF REGRESSION FREQUENCY AS A MEASURE OF READING PERFORMANCE

	Easy Prose				N = 77				
	First Sitting				Second Sitting				A 11
Selection	A	B	C	A–C	D	E	F	D–F	A–F
P	.71	.66	.58	.71	.75	.50	.62	.72	.78
A		.76	.64	.85	.69	.58	.65	.72	.81
B			.67	.90	.75	.60	.63	.75	.86
C				.84	.69	.65	.73	.76	.83
A–C					.77	.65	.71	.80	.93
D						.72	.74	.94	.91
E							.71	.86	.82
F								.88	.86
D–F									.96
	Hard Prose				N = 77				
	B	C	D	A–D	E	F	G	E–G	A–G
A	.66	.64	.70	.83	.65	.61	.65	.66	.79
B		.69	.67	.81	.75	.67	.65	.74	.87
C			.78	.90	.81	.70	.73	.79	.86
D				.93	.71	.70	.74	.79	.86
A–D					.83	.76	.80	.86	.91
E						.81	.81	.93	.89
F							.76	.88	.86
G								.94	.91
E–G									.94

tends to be slightly more typical of performance in general. These trends are seen in both easy and hard prose. Similar trends are discovered in Tables 5, 6, and 7 which contain the reliability coefficients for fixation frequency, regression frequency and pause duration. Practice in reading before the camera brings, therefore, only a slight improvement in consistency of performance. This change may be due to adaptation to either the laboratory situation, or to the reading material, or both.

As one would expect, test-retest reliability for material read at the same sitting yielded slightly higher coefficients

than the first sitting versus the second. For one paragraph versus others read on the same day the average reliability coefficient for perception time was on the average approximately .76 for easy prose and .74 for hard prose. When performance on the first day is correlated against that on the second, the average coefficients were .72 and .60 respectively for easy and hard prose. Similar trends appear in Tables 5, 6, and 7 for the other eye-movement measures. The range of these coefficients is .65 to .81.

TABLE 7

RELIABILITY OF PAUSE DURATION AS A MEASURE OF READING PERFORMANCE

	Easy Prose				$N = 77$				
	First Sitting				Second Sitting				
Selection	A	B	C	A–C	D	E	F	D–F	A–F
P	.68	.78	.78	.87	.60	.67	.65	.61	.80
A		.77	.73	.85	.68	.68	.66	.70	.82
B			.89	.96	.74	.77	.76	.79	.93
C				.94	.76	.77	.76	.78	.93
A–C					.79	.80	.80	.82	.95
D						.77	.85	.93	.89
E							.88	.88	.87
F								.93	.90
D–F									.94
	Hard Prose				$N = 77$				
	B	C	D	A–D	E	F	G	E–G	A–G
A	.82	.84	.73	.91	.71	.71	.75	.76	.85
B		.84	.76	.89	.74	.67	.76	.75	.87
C			.84	.95	.77	.68	.77	.78	.91
D				.90	.78	.73	.83	.82	.90
A–D					.83	.75	.85	.85	.96
E						.80	.83	.94	.90
F							.82	.90	.83
G								.98	.89
E–G									.94

The coefficients cited indicate that single short paragraphs of reading yield eye-movement measures that are adequate for group comparisons. To obtain coefficients which indicate that the reading measure is satisfactory for individual diagnosis it is necessary to read more material. When paragraphs A–C of the easy material are correlated with D–F (23 versus

23 lines), and A–D with E–G (39 versus 38 lines) of the difficult prose, an average coefficient of .82 is obtained. Only one coefficient, .72 for perception time on hard prose, is low (Tables 4–7). If individual diagnosis is planned, therefore, about 25 lines of material should be read. Results from less than approximately 20 lines probably should never be used for individual diagnosis unless a reliability of .80 or above has been established for the particular selection to be employed. While 10 to 15 lines sometimes yields a reliability of .80 or above, this is far from common.

It is interesting to note in Tables 4–7 that even with a very small sample of 5 or 6 lines of reading, as selection A versus F in easy prose and B versus F in difficult prose, the eye-movements are remarkably consistent from paragraph to paragraph (r = about .70). Doubling the number of lines raises the coefficients somewhat, but not as high as predicted by the Brown-Spearman formula.

Consistency of performance is little affected by variation in difficulty of material when a single sitting is considered. For one sitting versus a second, however, perception time, fixation frequency and pause duration are slightly more reliable and regression frequency slightly less reliable in easy than in difficult reading. The increased reliability of regression frequency in the difficult reading is undoubtedly due to the fact that relatively many more regressions are made in reading this material.

Fixation frequency, in general, had about the same reliabilities as perception time. Regression frequency was slightly less reliable than perception time and fixation frequency, and pause duration slightly more reliable than the same measures.

Our results indicate that eye-movement records have satisfactory reliability as measures of reading performance. This is true for even very small samples (5 lines) where group comparisons are to be made. It should be noted that our reliability coefficients are smaller than those obtained by odd-even correlations in previous investigations.

Validity

In the customary set-up for photographing eye movements a strong light shines into the reader's eye and the head is firmly clasped in a headrest. In addition, the whole experimental room usually presents a decidedly strange environment. Is the subject able to give a typical reading performance under these conditions? It is true that some of the subjects complain of the bright light and others appear a little bewildered at first by the unfamiliar situation.

To study the adaptation of the subjects to the laboratory situation, reading in the regular manner at a table was compared with reading before the camera as described in procedure. Form A of the Chapman-Cook test was read in the regular manner, and Form B before the camera. Form B is equivalent to Form A when Form B is given second and the testing conditions are identical. The scoring was the same in both situations. Any effect of the camera situation should appear as a difference between the two scores. The results are given in Table 8. The comparison of scores shows that, on the average, the subjects read as well before the

TABLE 8

Showing Adaptation to the Laboratory Situation. $N = 57$ Freshmen

Test and Condition	Mean	σ_M	r: Avs. B	Diff.	$\frac{D}{\sigma_{Diff.}}$
CC Form A: Regular.....	16.08	.51	.94	.05	0.28
CC Form B: At camera.....	16.13	.55			

camera as when reading in the customary manner. Although some subjects did better and some poorer before the camera, the group as a whole gave an entirely typical performance in the photographic situation.

These data do not mean that no attempt should be made to help the subject adjust to the artificial situation. As noted in our procedure, care was taken to explain the operation of the apparatus and a practice trial was given before the main experimental trial. This always tends to remove much of

the strangeness from the situation and to calm those who show nervousness.

As a check on previous findings the various eye-movement measures were correlated with achievement on standardized reading tests as criteria. The results for fixation frequency versus the tests are given in Table 9. The validity coefficients are quite comparable to Litterer's findings (cited above).

TABLE 9

VALIDITY OF FIXATION FREQUENCY AS A MEASURE OF READING. CRITERIA = STANDARDIZED READING TESTS

Text	E–M Selection	Performance Tests					
		CC Speed		I. Comprehension		M.R.E. Vocabulary	
		r	Att. *r*	*r*	Att. *r*	*r*	Att. *r*
Easy Prose	P	−.32	−.41	−.18	−.22	−.40	−.47
	A	−.49	−.61	−.35	−.41	−.35	−.40
	A–C	−.49	−.59	−.33	−.38	−.38	−.43
	D–F	−.49	−.60	−.36	−.42	−.46	−.51
	A–F	−.49	−.58	−.33	−.37	−.42	−.46
Hard Prose	A	−.28	−.38	−.17	−.22	−.27	−.33
	B	−.47	−.63	−.24	−.30	−.47	−.57
	A–D	−.44	−.58	−.32	−.40	−.40	−.47
	E–G	−.41	−.54	−.28	−.34	−.38	−.45
	A–G	−.40	−.50	−.31	−.36	−.38	−.43

The range is − .22 to − .61 (attenuated *r*'s) for easy prose and − .22 to − .63 for hard prose. Tables will not be given for the other measures which showed similar trends. The ranges of the validity coefficients are: (1) perception time: − .30 to − .64 for easy prose and − .24 to − .66 for hard prose; (2) regression frequency: − .06 to − .39 for easy prose and − .08 to − .41 for hard prose; (3) pause duration: + .02 to − .25 for easy prose and + .01 to − .23 for hard prose.

Certain trends should be noted in these coefficients: (1) Regression frequency and pause duration are much less valid than either perception time or fixation frequency. (2) In practically all cases the first selection read was less valid than subsequent paragraphs. Apparently adaptation to the reading situation (material and method) yields more valid performance. (3) None of the validity coefficients are large.

Medians for the four measures are fixations, − .43; perception time, − .44; regression frequency, − .28; and pause duration, − .17. Keep in mind that the material read before the camera was different than that in the performance tests. Other reports [11] have shown that there are many specific reading skills rather than a general reading ability, and that the intercorrelations (one test score vs. another) to be expected range from .00 to about .60. The above cited coefficients, therefore, probably represent good validity. At the best, however, conclusions based upon results derived in this manner are hazardous.

A more satisfactory way of obtaining validity of eye-movement measures is to read before the camera material which is strictly comparable to that in the test employed as a criterion, and to maintain before the camera and on the reading test the same reading attitude. As explained in the procedure, part of the Chapman-Cook Speed of Reading Test was read before the camera and eye movements photographed. Comparison of these records with the performance scores on the test should yield a satisfactory method of computing validity.

TABLE 10

VALIDITY OF EYE-MOVEMENT MEASURES WITH PERFORMANCE ON COMPARABLE READING MATERIAL AS CRITERION

CC Performance Compared with	r	Att. r
CC_1 Fixations	−.69	−.99
CC_2 Fixations	−.56	−.80
$CC_1 + CC_2$ Fixations	−.71	−.93
CC_1 Perception Time	−.65	−.88
CC_2 Perception Time	−.64	−.87
$CC_1 + CC_2$ Perception Time	−.71	−.90
CC_1 Pause Duration	−.08	−.10
CC_2 Pause Duration	−.24	−.31
$CC_1 + CC_2$ Pause Duration	−.14	−.17

The validity coefficients thus derived are shown in Table 10. The correlations, when corrected for unreliability, are very high for fixation frequency and perception time, and low for pause duration.[12] The average validity coefficient for

[11] Tinker, M. A., *Psychol. Rev.*, *op. cit.*

[12] These high coefficients are confirmed in an unpublished study done at the Minnesota Laboratory. Because of the nature of the reading material in this part of the experiment, regression frequency could not be computed with a satisfactory degree of accuracy and was therefore omitted.

the first two measures is — .895. It is not surprising that the coefficients are so low for pause duration since we know that this is not an adequate measure of reading speed.[13] Had we been able to compute the validity of regression frequency it would probably have been only a medium sized coefficient for it is only a fair measure of reading speed. Fixation frequency and perception time, therefore, are highly valid measures of reading performance.

In contrast with the high correlations (for fixation frequency and perception time) in Table 10, we have the correlations of the eye-movement measures in reading the Chapman-Cook Test material with performance on the Iowa Comprehension and on the Minnesota Vocabulary. They range from — .41 to — .69 with a median of — .54. Again the coefficients are relatively very low for pause duration: .00 to — .38 with a median of — .11. The intercorrelations for the three performance tests are .43, .46, and .49. Thus we find again only correlations of moderate size when variation in reading material occurs from one reading situation to another, and again little correlation of pause duration with reading speed or comprehension.

Summary and Conclusions

1. The reliability and validity of eye-movement measures were investigated for reading materials varying in length and in difficulty.

2. The main group of readers consisted of 77 university sophomores. A sub-group contained 57 freshmen.

3. The eye-movement measures, even for reading selections of 5 or 6 lines, have adequate reliability where group comparisons are concerned. This is true for even the first selection read before the camera.

4. Reliability increases slightly with practice in reading before the camera.

5. When individual diagnosis is planned, adequate reliability is ordinarily achieved only if 20 or more lines of material are read.

[13] Tinker, M. A., *Psychol. Rev., op. cit.*

6. Reading before the camera yields typical reading performance. To insure adequate adaptation to the laboratory situation, however, a practice trial in reading before the camera should be given.

7. Comparison of eye-movement measures with performance test scores, when the material read before the camera is different from that in the tests, is an unsatisfactory method of computing validity of the measures.

8. When the material read before the camera is strictly comparable to that in the performance tests, the validity of fixation frequency and perception time is very high.

(Manuscript received November 25, 1935)

THE ILLUSION OF THE OBLIQUE INTERCEPT

BY PAUL SCHILDER AND DAVID WECHSLER

From the Psychological Department of the Psychiatric Division of Bellevue Hospital, New York University Medical College

The literature of optic illusions is so enormous and the interpretations given them are so numerous that one hesitates to add to them. Our justification for doing so now is, first, that the illusion to be presented has, so far as we know, not been described before, at least not in its present form; second, that the experiments to be detailed seem to throw some additional light on the laws of form perception.

The optic illusion to which we wish to call attention and which we shall call the *illusion of the oblique intercept*, consists of a number of peculiar optical phenomena which are produced in any system of narrowly spaced parallel lines when cut by a transversal. The effects produced are in both parallel lines (pr.l.) [1] and intervening spaces (i.sp.) [1] between them and are to be observed at the edges of intersection. The phenomena are illustrated in Fig. 1. Here a series of lines of approxi-

FIG. 1.

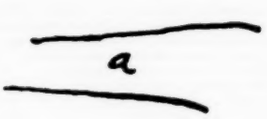
mately 1 mm in width are drawn on a white background at a constant distance of about 1½ mm with a strip of white paper placed across the lines so as to cut them at an angle of about 45° from the vertical. Fixating on the edges of intersection

[1] In subsequent references to the figures, pr.l. stands for parallel lines; i.sp. for the intervening spaces; and trs. for the cut transversal or intercept.

one will observe (1) that the horizontal lines at the point where they are cut by the transversal are slightly clubbed; (2) that the spaces between them seemingly diverge; (3) that where the transversal meets the lines they appear to curve; (4) that the curvature or bending is downward—that is, obliquely towards the horizontal; (5) that in the case of any given line the curvature of its upper edge is always less than that of the lower edge, so that its cut thickness viewed singly appears to be arched; (6) the intervening spaces (i.sp.) show the same phenomena as the parallel lines (pr.l.). It is immaterial whether the i.sp. or pr.l. are seen as figure or background; nor does it matter whether the trs. is a white or black line (see Fig. 2).

FIG. 2.

A pair of lines showing the phenomena just described in a grossly exaggerated fashion is shown in Figs. 4a and b, in order to illustrate the features which may escape the unpracticed observer.



FIGS. 4a and 4b.

There are a number of factors which influence the intensity of the illusion. The most obvious of these is the angle at which the oblique transversal crosses the parallel lines. If the transversal is placed perpendicular to the lines, that is at an

angle of 90°, the illusion is entirely absent. The clubbing of the horizontal lines, for example, does not become noticeable until the transversal is rotated some 15° from the vertical (that is, forms an angle of about 105° with the intersected horizontal). From that point down it increases progressively until an angle of about 140° to 150° is reached, after which the phenomena remain fairly constant for another 15° of clockwise rotation. As an angle of about 165°, however, is reached, some rather peculiar phenomena take place. The clubbing of the lines persists, but there is a change in the direction of their curvature; instead of both edges of the intersected pr.l. curving downwards, the upper edge now seems to be turned upwards; whereas the lower edge continues to curve in its original direction (Figs. 6a and 6b).

If the transversal instead of being rotated from the perpendicular towards the horizontal, is placed parallel to the horizontal and then rotated from 0 to 90° (counter clockwise), similar phenomena will be observed except that the effects described at the critical angles will be reversed.

In addition to the foregoing there are a number of other details regarding the phenomena which are well worth noting. Among the most important are the following:

It does not matter whether the parallel lines are horizontal or vertical. Of importance is only the angle at which the transversal cuts them.

The illusion is less striking if the parallel lines are very thin or very thick. (Under ordinary conditions, a maximal effect is obtained with parallel lines having a cross-section of about 1 mm; spaced at a distance of about $1\frac{1}{2}$ mm from each other.) If the lines are too thick, the curvature appears to be only slight, and the clubbing of the lines tends to disappear. The latter, probably, because the illusory effect is relatively negligible when the diameter of the lines is already considerable.

Darkening of the visual field has no measurable effect on the phenomena.

When the parallel lines are short, the phenomenon decreases considerably. Some extension of the pr.l. is seemingly

necessary in order to give the impression of a difference between the oblique and perpendicular cut of the transversal. This difference is particularly outstanding when one has a number of pr.l., and covers them obliquely with a black paper so that the lowest line has about a length of 20 mm, the subsequent line about 25 mm, etc. (See two lowest lines in Fig. 2.)

The phenomenon of bending and clubbing can be seen even when only one parallel line is used, but it is much more marked when we deal with a series of lines. This is true even when the distance (i.sp.) between the parallel lines is increased to about 1 cm; but it is more outstanding when the distance between the horizontal lines is only about $1\frac{1}{2}$ or 2 mm.

The illusion persists under tachistoscopic exposure down to 1/25 of a second, though under short exposures it is greatly attenuated; generally speaking, the shorter the exposure, the less striking is the perceived illusion.

Fatigue and fixation of long duration accentuates the illusion.

We were not able to produce the illusion stereoscopically. It was not possible to combine corresponding figures owing to visual rivalry.

The illusion is particularly marked if the lines are observed at a distance sufficient to give one a total impression of the system. Under this condition the area about the parallel lines in the region of their interception seems to be extended, and both the lines and the intervening spaces seem to diverge like the strings at the base of a harp.

There is also a marked alteration in the appearance of the cutting transversal (trs.). In the region where the clubbing occurs, whether on the black lines (pr.l.) or on the white spaces between (i.sp.), the intersected portion of the transversal seems to be definitely curved inwards. The intercept (trs.) as a whole is also slightly curved, and the entire transversal takes on convex appearance, encroaching as it were on the spaces between the parallel lines (pr.l.). This effect is particularly striking if one attends to the figure as a whole. (It is as if the whole trs. system has undergone changes

similar to those of the intercepted parts of the single pr.l. and the spaces between them.) (Figs. 1–4.)

If one concentrates on the appearance of the cutting line (trs.), one finds that it does not appear as a straight line but looks like a staircase, the parallel lines representing the succeeding levels or 'steps' of the staircase. The transversal is broken into small pieces each succeeding one of which is pushed inwards as it approaches the next higher 'step' of the pr.l. system. This staircase phenomenon (Fig. 3)

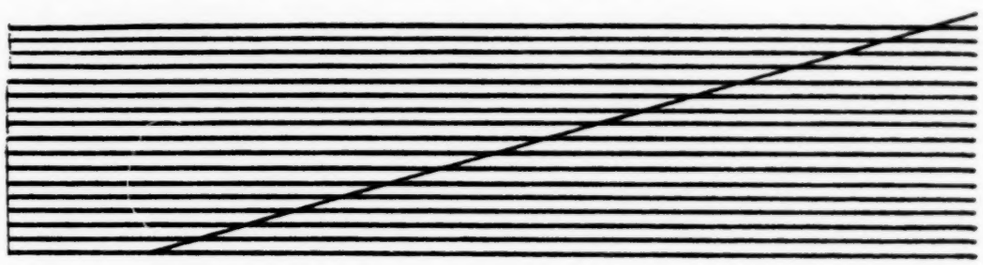


FIG. 3.

augments as the cutting trs. is rotated from the vertical towards the horizontal. It is particularly clear as it approximates an angle of 10 degrees from the horizontal, where the illusion of a staircase is almost complete (Figs. 6a and b).

As may be supposed, the explanation of the illusion is more difficult than its description. The interpretation of geometrical illusions generally call upon three principles: There is, first, the illusion of perspective, or the tendency of lines to take on three dimensional aspects, as in the pseudoscopic prisms or the Oppel figures. Second, there is the illusion of variable space, where different portions of space are attracted to the figure, depending upon the orientation of the limiting lines, as in the case of the Müller-Lyer illusion. Third, there is the illusion resulting from the tendency to overestimate small angles and underestimate large angles, as in the case of the Poggendorf phenomenon.[2]

[2] To these are often added some subsidiary principles, like the tendency to see obtuse angles as right angles, the attraction of the horizontal for the oblique at its intersection, the influence of eye movement in traversing a figure, etc. But none of these taken alone seem sufficient as ultimate explanations. We agree with Koffka that these principles take on significance only when considered in connection with more fundamental *Gestalt* principles.

Of these three, it is clear almost at once that the first two, namely the principles of perspective and of variable space, can be of little assistance to us. On the other hand, the principle of over and under estimation of angles seems, at first, quite pertinent. This particularly, in view of the fact that the illusion of the oblique intercept has a rather close resemblance to the Poggendorf phenomenon, which is commonly accounted for in terms of the overrating of sharp angles. The resemblance between the illusion of the oblique intercept, or at least certain aspects of it, to that of the Poggendorf phenomenon is perhaps best seen from the simultaneous inspection of the modified Poggendorf figure as illustrated by Wundt (Fig. 5) in his '*Physiologische Psychologie*.'[3] Wundt describes

FIG. 5.

it as follows: "A vertical black stripe is crossed at a sharp angle by an oblique line. The ends of the oblique line bordering the stripe seem to be bent towards the stripe."[4] The oblique intercept (trs.) in our illusion corresponds to the crossing line in the Poggendorf phenomenon. The Poggendorf phenomenon, however, never has been sufficiently studied with any considerable number of the striped lines arranged into a system, and the phenomena described by us have,

[3] 6th edition, Vol. 2, page 588:

[4] Werner has further noted that the vertical stripe is also curved.

therefore, escaped the attention of previous investigators. Although one observes in our figures the continuity of the pr.l. system as one does in the crossing line of the Poggendorf phenomenon, we nevertheless feel justified in attempting an interpretation of them based upon the phenomena primarily observed in our illusion.

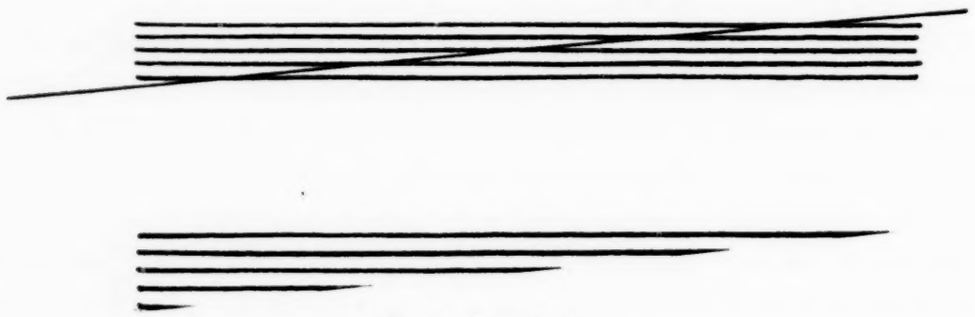
Of fundamental importance in the interpretation of the illusion of the oblique intercept is the fact that the horizontal lines are curved in the region where they are intersected by the transversal. One cannot speak, therefore, merely of the overrating of small angles, since we are here confronted with a new phenomenon, namely a curving effect. This curving may be observed in Figs. 1 and 2 and is particularly noticeable on the lower part of the parallel lines. No interpretation can neglect this curving[5] which cannot be referred back to any other phenomenon. It is not dependent on the luminosity of the crossing system (trs.) nor is it due to irradiation. The tendency to curving on the other hand is closely bound up with the characteristic of the crossing line (trs.). Thus, it increases with the diminution of the angle between the parallel system (pr.l.) and the intercept (trs.). Our view is that the mass of the parallel lines attracts the mass of the crossing system as a whole, and also the single portions of the crossing system. The curving of the parallel lines is therefore less noticeable when the crossing line (Fig. 3) is thin and accordingly exerts less mass attraction. If such a mass attraction takes place one would expect it to be stronger on the lower edges of the pr.l. which are more exposed to the attraction of the crossing system, owing to the reduced inner pull of the pr.l., *per se*.

It is very probable that the tendency to see angles as curves is to a great extent due to the mutual attraction between their two sides. Any two intersecting lines seemingly exert a mutual pull on each other, and the phenomenon with its consequent effects (curving) is not restricted to lines cutting each other at sharp angles. Thus F. B. Hoffman

[5] In describing the modified Poggendorf phenomenon, Wundt refers to the bending of the cutting line. This is quite different from a curving, and is an incorrect description of what actually occurs.

has shown that even very obtuse angles appear curved when the wides of the angle are made very small.

When the angle between trs. and pr.l. becomes 165 degrees, the upper edge of pr.l. is curved upwards as mentioned above (Fig. 6). At this angle the mass attraction of the trs. be-



FIGS. 6a and 6b.

comes paramount, also for the upper edge of pr.l., whereas the inner pull of the pr.l. is reduced. It thus seems that the tendency to see lines curved in virtue of their mutual mass attraction is a widely recurring phenomenon. The attraction between lines is dependent on the mass of the figure and on the relation of the different parts to one another. Many details of our findings point into this direction. The mass distribution in figures plays the most important part. Our view is that the overrating of small angles in our figures is due to the curving of the enclosing lines at the point where they meet. Contrary to Wundt's opinion, there is no suggestion that either perspective or overrating of such angles is the cause of the curving.

The clubbing of the lines may be interpreted as being due to different attractions exerted on the upper and lower borders of the pr.l. stripes by the trs., though it is very probable that this explanation is insufficient. Another factor to be considered is that when a transversal cuts a stripe (pr.l. or i.sp.) at an angle, the diameter of the section is actually much greater than it is when cut perpendicularly. We tend to orient ourselves towards the vertical; this, seemingly because we are accustomed to see lines cut at approximately right angles. The tendency is very general and was already noted by Fechner and used by him in explaining "Why sausages are

cut on the slant." The answer is that the slice looks much bigger. The sausage is more imposing, since the obliqueness of the cut has not been taken into full consideration. That this effect plays really a part in the phenomenon, is reinforced by the illusion experienced from the rotation of the transversal through an angle of 90°. During such rotation each segment of the cut transversal changes from the vertical to the horizontal. In doing so it broadens out, and this change is closely related to the clubbing effect. The effect produced by the gradual rotation of the transversal suggests that the mutual mass attraction is particularly strong when there is movement in the optic field, and also that intrafigural movement may be an important factor in what we have called mass attraction. There is no evidence that eye movements, *per se*, play an important part in this phenomenon.

We are inclined to underrate the deviations from the fundamental directions of orientation, namely the horizontal and the vertical. When the horizontal line is intercepted obliquely the deviation of the crossing line from the perpendicular will be reconstituted, that is, will tend to appear nearer to the vertical than it really is. When we see the intercept nearer to the vertical, the effect of the apparent rotation is to make the intersected bit of the intercept appear to be thicker.

It is especially difficult to explain the changes in the crossing system (trs.). They seem to be closely related, however, to the curving-in phenomenon which is present both in the parallel lines and the intervening spaces. In any case, there is a seemingly stronger attraction between the trs. and the lower border of the intercepted lines, than between the trs. and its upper border. This difference is especially marked when the transversal cuts the parallel lines at a very small angle to produce the staircase phenomenon already mentioned (Figs. 3 and 6).

The curving-in phenomenon is the expression of the effect of the counter mass attraction of the parallel lines on the transversal. As might be expected, the smaller the internal angle between the two intersecting lines the greater will this

effect be. Under such conditions, particularly when the angle is very small, if we focus our attention on the transversal segment between any two parallels, it will be observed that the lower edge of the upper parallel exerts a marked attraction for the upper third of the transversal segment, and the upper edge of the lower parallel a similar attraction for the lower third of the transversal segment. The middle third appears to be brought to a vertical position so that the total illusion is as shown in Fig. 7. Here again then, the observed phe-

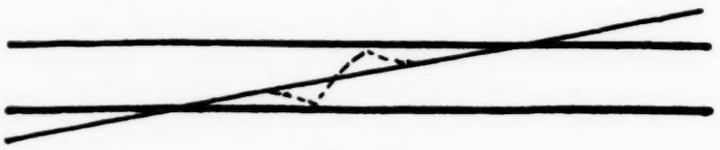


FIG. 7.

nomenon seems to be mainly due to two factors, (1) mass attraction, (2) the tendency to orient towards the vertical or the horizontal.

In the final interpretation of the illusion, therefore, we are left chiefly with the problem of mass attraction between the different parts. This mass attraction might be referred to as a sort of 'gravitational pull.' This principle is bound up with the principles of intrafigural movements. Our view is that mass attraction is the basic principle which explains the phenomena we have described. Supplementing it are the principles of intrafigural movements and the tendency of the human eye to orient all intersecting parts towards its preferred directions, namely the horizontal or vertical. All the other principles such as the overrating of angles, the tendency to see curves instead of straight lines, are seemingly subordinated to these two. The principle of mass attraction has to do with the organization of the total visual field, and thus explains the enormous influence which the repetition of the motive has on the illusion. We think that the principles here presented will also furnish a better explanation of the Poggendorf phenomenon, and the Herring and the Oppel figures for which Sanford ventured a similar interpretation.

SUMMARY

The authors describe an hitherto unreported optical illusion, with the factors which seemingly determine it. The

illusion, termed *the illusion of oblique-intercept*, is observed whenever a system of parallel lines is crossed by a line cutting them obliquely. The main characteristics of the illusion are a clubbing of the parallel lines at the edge of interception, an apparent spreading or 'fanning' of the intermediate spaces, a curving of these as well as of the intercepted line, and finally a staircase effect along the intercepting transversal.

The phenomena described are accounted for in terms of a number of psycho-physiological tendencies of which the two principal ones are: 1. The phenomenon of mutual mass attraction of the sides of intersecting lines upon each other. This mutual attraction accounts for the rounding up of sharp angles and the overrating of small angles. The mutual mass attraction is closely related to the total organization of the visual field and its intrafigural movements, and increases, therefore, with reduplication. 2. The human eye has a preference for the vertical and horizontal, and accordingly shows a tendency to orient points of the compass to one or another of these principle directions. In virtue of this preference there is a further tendency to neglect directional deviations from the principal visual coördinates. In consequence of this neglect, the sides of the angles depending upon characteristics of the visual field, are attracted nearer to the one or the other of the main visual coördinates.

This last phenomenon is particularly clear in the case of the staircase illusion in which the deviations of an oblique line are in part attracted to the horizontal and in part to the vertical directions of the crossing system.

(Manuscript received November 11, 1935)

References

1. Hoffman, F. B., Die Lehre vom Raumsinn, 1st and 2nd edit. *Hand. der Gesamten Augenheilkunde* by Axenfeld and Elschnig, 2nd edit. Vol. 3, Chapter 13, Berlin, Springer, 1925.
2. Koffka, K., *Psychologie der optischen Wahrnehmung. Ibid.*, pp. 1215–1271. *Handb. d. norm. und path. Physiol.*, 12, 2.
3. Sanford, E. C., *Course in Experimental Psychology.* 1898.
4. Werner, H., Uber Structur Gezetze in der sogenanten Geometrich optichen Tauschungen, *Zeit. f. Psych.*, 1929, 14, 248.
5. Wundt, W., *Gundzüge der Physiologischen Psychologie.* Vol. 2. 6th edition, 1910.

MAZE LEARNING WITH A DIFFERENTIAL PROPRIOCEPTIVE CUE *

BY LAUNOR F. CARTER

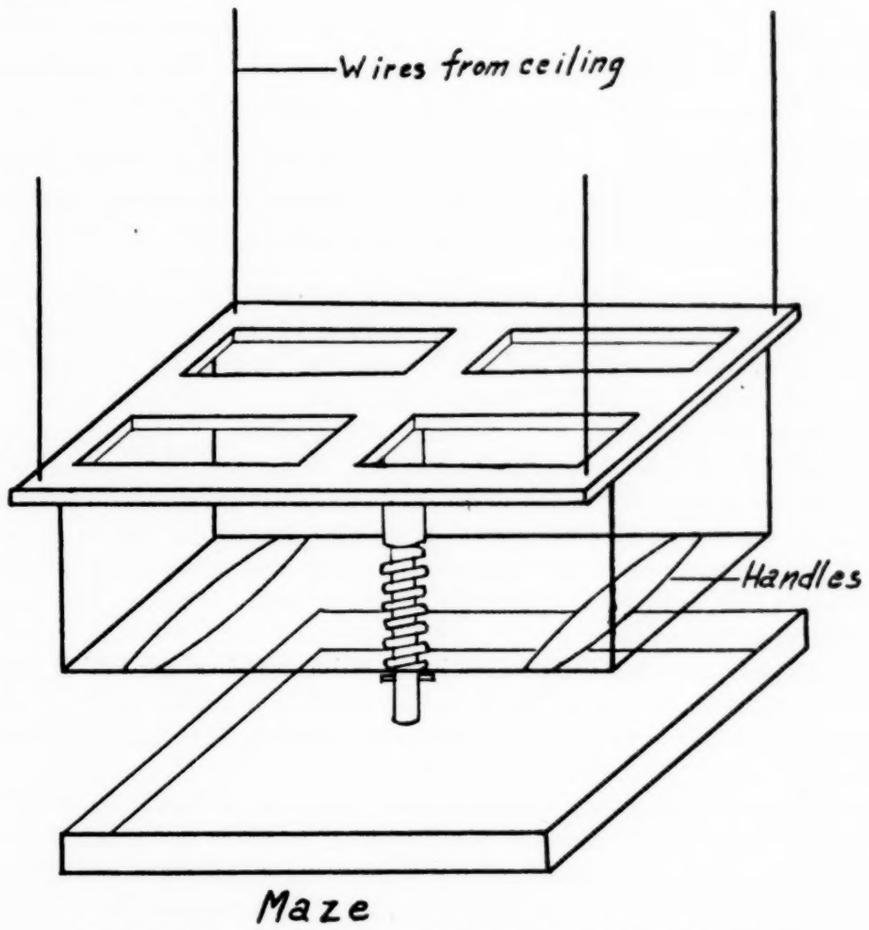
The effect of specially reserved tension patterns produced by definite muscular contraction has not, in some recent experimentation, been sharply differentiated from general tension. Specific proprioceptive patterns may act as cues for specific responses. The results presented in this paper show that learning which uses specially reserved tension patterns as a cue for conditioning is definitely more efficient than comparable learning without such a cue.

The role of the proprioceptors in conditioning was first stressed by Guthrie in 1930 (4). Recently (1935) (5) he has reiterated his belief in the importance of this type of stimulation. When Guthrie first pointed out the role of the proprioceptors in conditioning, there was little direct experimental work to support such a view even though the hypothesis was reasonable enough. Bills (1) had reported increased efficiency in learning when tension was artificially introduced by a hand dynamometer. Jacobsen (6) had reported his experimentation with progressive relaxation. Since then, however, Stroud (10), Freeman (3), Chase (2), McTeer (7), and Zartman and Cason (11) have reported experiments bearing indirectly on the problem. Among these experimenters there is considerable disagreement as to the effect of general tension on learning. However, there is general agreement that some minimal amount of muscular tension is probably necessary for efficient learning. In 1932 Sheldon (8) devised an experiment which was intended to test Guthrie's hypothesis. Her results were largely negative; however, Smith (9) has recently criticized the experiment both on

* The author wishes to acknowledge the kind assistance of Dr. Stevenson Smith throughout the experimental work and the preparation of the manuscript. He also wishes to acknowledge the assistance of V. D. Bain for running a number of the subjects.

theoretical and experimental grounds. In the same paper Smith reported an experiment which gave some experimental support to Guthrie's hypothesis.

The present experiment is an attempt to test Guthrie's hypothesis on empirical grounds. To test the hypothesis adaquately it is necessary to devise a learning problem to which can be added, at will, a differential proprioceptive cue. The maze problem was selected as most adequately satisfying this need.



View of stylus and maze in position to run without the springs.

Three features unusual to maze construction were used in making the stylus and mazes. First, the friction between the stylus and the maze was kept constant irrespective of the downward pressure exerted by the subjects. This was accomplished by a wire suspension of the complete stylus. Second, the subjects were not blindfolded. The mazes were screened by being immersed in a pan of oil. Third, the

stylus was constructed so that springs could be attached to it. These springs served as a definite proprioceptive stimulus. One maze was learned with the springs on the stylus and another without the springs by each subject.

When the springs were attached to the stylus, the subjects had to exert a muscular force to overcome the tension offered by the springs. At every point in the maze the subjects had to exert a different muscular force to overcome this tension. At times this muscular response was one of magnitude and at others one of direction. This change of muscular exertion required at all points in the maze undoubtedly stimulated a different proprioceptive pattern at every point in the maze. If this assumption is correct and Guthrie's hypothesis is veritable we should expect more efficient learning with the springs connected to the stylus than without.

Two mazes were constructed which shall be designated as A and B throughout this paper. Mazes A and B were identical except as to pathway. The pattern used was the same as that reported by Smith (9) except that the last unit of the first line of A reads R(F) instead of R(FL). Onto a 12″ × 12″ × ½″ piece of fiberboard ¼″ plyboard strips were bolted to form the sides of the paths. The pathways were ¼″ wide; the mazes being constructed in 2″ units. At the beginning and end of each unit a brass plug was inserted in the floor of the maze and ran through the fiberboard to a brass plate on the underside of the maze. These plugs were flush with the floor of the maze. When the maze was being traversed the stylus made contact with these plugs which caused a light to flash on. This light always indicated to the subjects that they had reached a decision point.

This complete maze was then bolted into a galvanized iron pan which was filled with oil to which enough carbon black had been added to make the mazes entirely invisible. Due to this screening of the mazes the subjects were not blindfolded and could use their ordinary eye-hand coördination.

The stylus was constructed by attaching a ⅝″ × 4″ brass rod to a 18″ × 11″ × 1″ cedar board. This rod was drilled to admit a rod ⅜″ × 7¾″. 6½″ of this ⅜″ rod was turned down to .225″. This small end was drilled and a ball fitted into the socket. One side of the ball had built on it a flat plane, .225″ in diameter. By this ball and socket arrangement a constant contact was kept with the floor of the maze. Handles were suspended from the board, about two inches clearance with the oil being allowed. The subject spread his hands, palms downwards, on these handles and conducted the stylus through the maze.

A 4″ coil expansion spring was slipped over the ⅜″ rod. This spring caused the stylus to maintain a constant tension on the floor of the maze irrespective of the downward pressure the subjects exerted on the stylus. This effect was accomplished by swinging the complete stylus from the ceiling of the experimental room. From each corner of the board, to which the stylus was attached, wires were run to the ceiling. This suspension allowed the stylus to swing in a great 93″ arc which kept the stylus at all times perpendicular to the floor of the maze, and provided an opportunity to supply

a differential proprioceptive cue. In different parts of the room four large coil springs were anchored. These springs could be attached to the stylus by hooks on the board. The springs were so arranged that when the stylus was at rest it was held in the center of the maze.

Procedure.—Forty subjects, largely from elementary psychology classes, were used in the experiment. Twenty of these were women and twenty men.

Two mazes were used in the experiment. Maze A showed a reliability of .802 which becomes .891 by the Brown formula for halved scores; maze B showed a reliability of .883 which becomes .937. These coefficients were calculated by the scoring of the odd against the even runs omitting the first run. Each subject was required to run one maze with the springs and the other without springs. Ten subjects, five men and five women, ran each of the following sets: pattern B run first with springs, pattern A run second without springs; pattern B run first without springs, pattern A run second with springs; pattern A run first with springs, pattern B run second without the springs; pattern A run first without springs, pattern B run second with springs. Each subject ran through each maze ten times, at whatever speed he chose. An error count was kept; counting errors enteries into the blind alley far enough to cause the light to flash and counting backtracking. Subjects were allowed to backtrack only one unit along the true pathway. The subjects were given a printed sheet of instructions as to the method of running mazes, purpose of the light, etc. They were then allowed to practice on a small practice maze. This maze was practiced with the springs on the stylus, if the first maze they were to run was with the springs, otherwise the springs were not used for practice.

RESULTS

The objective results show that the added differential proprioceptive cue produced definitely superior learning to learning without this immediate cue.

TABLE 1

SCORE CORRECTION FOR ORDER OF PRACTICE AND PATTERN DIFFICULTY

Maze	Total Errors	Correction Ratio
Practiced first	1886	1.00
Practiced second	1520	1.24
Pattern A	1855	.84
Pattern B	1551	1.00
Pattern A practiced second		1.04

TABLE 2

ERROR SCORES

Maze	Total	Correction Ratio	Corrected Totals
A1O	530	.84	445.3
A1W	588	.84	493.5
A2O	409	1.04	425.3
A2W	328	1.04	340.1
B1O	436	1.00	436.0
B1W	332	1.00	332.0
B2O	471	1.24	586.4
B2W	312	1.24	386.3

A and B for maze patterns. 1 and 2 for practiced first or second. O and W for practiced without or with the springs.

All scores are thus corrected to conform to the difficulty of pattern B practiced first.

TABLE 3

COMPARISON OF CORRECTED ERROR SCORES, WITH AND WITHOUT SPRINGS *

Maze	*N*	*M*	S. D.
Without	40	47.42	25.21
With	40	38.79	19.61

$r_{wo} = .696 \pm .055$. S. D. of Diff. = 2.87. Diff. = 8.63.

$$\frac{\text{Diff.}}{\text{S. D. of Diff.}} = 3.00. \quad p = .999.$$

* The S. D. of the Diff. computed from the raw scores, uncorrected for maze difficulty or order of presentation, is 2.13. The corresponding p being .983.

CONCLUSION

As can be seen from Table 3 there is a reliable certainty that the difference between the error scores with the springs and scores without the springs is significant. It would appear in consideration of the significant reliability of the difference in error scores, as obtained with and without differential proprioceptive cues, that this experiment adds empirical support to Guthrie's hypothesis.

(Manuscript received November 11, 1935)

REFERENCES

1. BILLS, A. G., The influence of muscular tension on the efficiency of mental work, *Amer. J. Psych.*, 1927, **38**, 226–251.
2. CHASE, W. P., The role of kinesthesis in ideational maze learning, *J. Exper. Psych.*, 1934, **17**, 424–438.
3. FREEMAN, G. L., The spread of neuro-muscular activity during mental work, *J. General Psych.*, 1931, **5**, 479–493.
4. GUTHRIE, E. R., Conditioning as a principle of learning, *Psychol. Rev.*, 1930, **37**, 412–428.
5. GUTHRIE, E. R. *The Psychology of Learning.* Harper. N. Y., 1935.
6. JACOBSEN, E. *Progressive Relaxation.* Chicago: Univ. of Chicago Press. 1929.
7. MCTEER, W., Changes in grip tension following electric shock in mirror tracing, *J. of Exper. Psych.*, 1933, **16**, 735–742.
8. SHELDON, M. I., The effect on learning and retention of an increase in the activity of proprioceptive sense organs, *Univ. of Calif. Pub. Psych.*, 1932, **46**, 296–302.
9. SMITH, S., The principle of specific conditioners, *J. Genetic Psych.*, 1935, **46**, 296–302.
10. STROUD, J. B., The role of muscular tension in stylus maze learning, *J. Exper. Psych.*, 1931, **15**, 606–631.
11. ZARTMAN, E., AND CASON, H., The influence of an increase in muscular tension on mental efficiency, *J. Exper. Psych.*, 1924, **17**, 671–679.

THE INFLUENCE OF DEGREE OF WHOLENESS ON WHOLE–PART LEARNING

BY M. V. SEAGOE

University of California at Los Angeles

Another article in this series will show that one cause of confusion regarding the whole-part problem is the failure to define the terms 'whole' and 'part' in a qualitative rather than a quantitative sense. The definition of whole here adopted is that of a Gestalt, that is, a definitely segregated, independent pattern which possesses unity, coherence, and meaning in itself above that implied by its parts. Conversely a part is an element in a total situation which is essential to the meaning of the whole, but which loses its peculiar meaning when isolated from the whole.

If it is the qualitative rather than the quantitative whole that determines the most economical unit of presentation, it should be possible to so arrange the situation that part learning (in the conventional sense) will be superior provided each part constitutes a stronger figure than the whole. Materials grading from loosely integrated figures to closely integrated ones were designed to test the hypothesis.

The materials consisted of the four block designs illustrated in Fig. 1. Designs were drawn in India ink on light cardboard, the cardboard glued to three-eighths inch plywood, and the plywood cut into sixteen pieces as indicated by the lines running through each figure. Problem 1 has four equivalent parts arranged symmetrically, yet without any unity other than that given by orientation and spatial juxtaposition. Problem 2 has the same elements plus some linkage, and problem 3 is a well-integrated form added to the same elements. Problem 4 is also well integrated, but the four segments are different rather than being the same motif turned in different ways; the result is less wholeness for each part. If the hypothesis that the economical unit will be that which most nearly constitutes a qualitative whole is correct, the 'part' method would be expected to prove definitely superior on the first problem, somewhat superior on the second, the 'whole' method slightly superior on the third, and definitely superior on the fourth.

The coefficients of reliability obtained from correlating odd against even trials for time scores on the whole method are presented in Table I. Results from the part method were eliminated because it was questionable whether combined scores would be

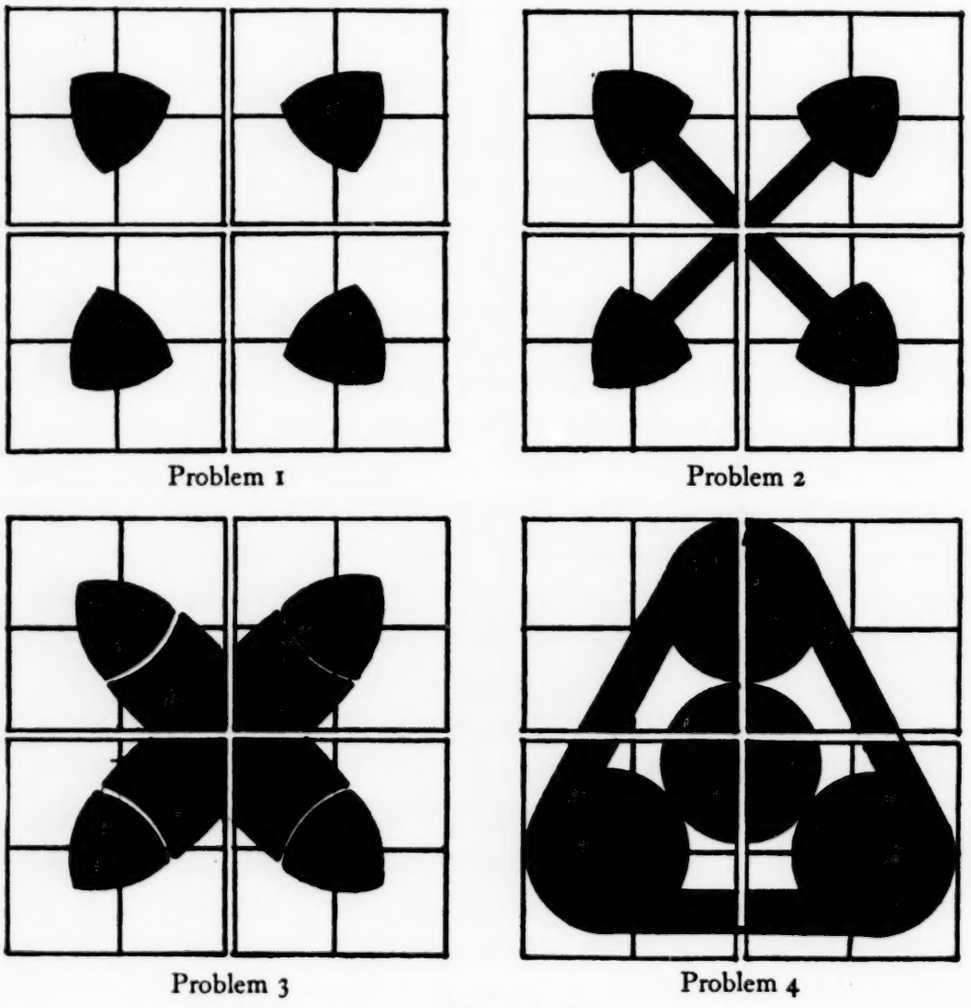


FIG. 1. Block designs.

significant in view of widely differing numbers of trials required for mastery. Results from the whole method where mastery was achieved at the first trial were also eliminated; in cases where there was an odd number of trials the first was disregarded. The population in each case is so small that it is questionable whether the use of a correlation

TABLE I

RELIABILITY OF MATERIALS

Problem	N	r	r by Spearman-Brown Formula
1	9	.351	.520 ± .164
2	8	.888	.941 ± .077
3	6	.435	.606 ± .174
4	10	.559	.717 ± .104

technique is justified. At any rate, the figures are suggestive and fairly consistent with the exception of Problem 2 which is undoubtedly spuriously high. The time involved in each problem was so short that no high reliability might be expected: the first took

about ten minutes, the second four minutes, the third two minutes, and the fourth five minutes as a rule. Further, since learning ceased at one correct reproduction, many of the coefficients are based on only two trials per subject, resulting in a lack of smoothing of chance fluctuations from combination. The problems seem reliable enough to justify their use in an experiment where the purpose is to compare two methods of using the same material. It is of interest to note that, with the exception of Problem 2, the reliabilities are in the order of degree of integration and complexity of the figure, though not necessarily in the order of time required for mastery.

The subjects were twenty-three students taking courses in psychology at Stanford University. One case was dropped because of failure to complete the series of problems. Subjects were assigned to Group I or Group II in order of arrival for experimentation. The results are shown in Table II. It is evident that the only great

TABLE II

CHARACTERISTICS OF SUBJECTS

Characteristic	Group I	Group II
1. Sex: Men	3	3
Women	8	8
2. Maturity: Graduates	2	1
Undergraduates	9	10
3. Thorndike Scores: Mean	78.11	84.80
Sigma	12.58	10.98

difference between the groups thus selected at random lay in the Thorndike Intelligence scores. The difference between the means is 6.69 points, the probable error of the difference is 3.67, the critical ratio is 1.82, and the chances of a reliable difference are 89 in 100.

On the basis of such evidence an intelligence factor might be argued in favor of Group II. The influence of such a factor was not constant, however, because of the plan of rotating presentation in such a way that each of the twenty-two complete cases learned two problems by each method. In addition, the coefficient of correlation between Thorndike score and speed of mastery for the series of four designs is $-.210 \pm .148$. The best assumption is that at the college level there is little or no relationship between intelligence as measured by the Thorndike test and efficiency in mastering block designs of the type used.

The method of presentation for odd-numbered subjects for designs 1, 2, 3, and 4 respectively was part, whole, whole, and part. For even numbered subjects the order was whole, part, part, and whole.

When the subject arrived he was given the following instruction: "I have four puzzles here. I will present two of them by a part method and two by the whole method. The idea is to see which of the two methods of presentation is better. Use just one hand, taking each block out of the box as you use it. Both the number of moves you make and the time you use will count on your score." Before each presentation the blocks were arranged in a standard shuffled position: for each quarter of the design the upper left and lower right blocks were interchanged as well as the upper right and lower left. No block was in its proper position, but the shuffling was so arranged that the position of presentation was the same in both whole and part presentation. A pattern showing the correct arrangement of the blocks on which the subject was working was before him while he copied the design. A screen prevented the subject's seeing the blocks or the pattern while the blocks were being shuffled.

When the subject began working the stop-watch was started, as he worked the

number of moves made was counted, and at the completion of the design the time record was taken. A move was defined as placing a block in relation to the design, whether or not it was the right position. Turning the block in the box or moving it out of the way was not counted; trying a block in a particular place in relation to those in the design, whether the hand left the block or not, was counted, as was also turning a block in place within the design. The definition of move is subjective and may not be valid in any absolute sense. It was consistent, however, from problem to problem and from method to method and hence does not influence the result of the experiment.

'Errors' were computed by subtracting the minimum number of possible moves from the actual number made. Time was recorded in seconds. Mastery consisted of one completion of the unit in the minimum number of moves, *i.e.* without error. Each unit was repeated until mastery was achieved, then the next presented and learned in a similar way.

With the part procedure part 1 was presented to mastery, followed by parts 3, 2, and 4 to a similar errorless reproduction. Parts were numbered clockwise beginning with the upper left corner. Then the whole was presented, and if not solved without error it was presented again and again until the criterion was reached. The whole procedure consisted of presenting all sixteen blocks and the total design each time from the beginning. Thus the part method used had some element of whole learning in it, diminishing the difference between the two methods.

TABLE III

RESULTS

Criterion	Problem			
	1	2	3	4
Hypothesis	P	P	W	W
Time: Mean, Whole Method	644.09	206.64	126.64	327.18
Part Method	608.82	167.09	156.27	453.73
P. E. Mean: Whole Method	94.93	32.77	11.43	34.10
Part Method	94.90	17.20	13.69	44.73
Difference Between Means	P 35.27	P 39.55	W 29.63	W 126.55
P. E. of Difference	134.23	36.98	17.83	56.25
Critical Ratio	P 0.26	P 1.07	W 1.61	W 2.27
Chances in 100 Reliable	P 57	P 76	W 86	W 94
Errors: Mean, Whole Method	19.91	4.36	1.36	9.73
Part Method	30.36	1.91	4.45	18.18
P. E. Mean: Whole Method	3.76	1.31	0.36	1.17
Part Method	8.51	0.37	1.34	3.54
Difference Between Means	W 10.45	P 2.45	W 3.09	W 8.45
P. E. of Difference	9.30	1.36	1.39	3.73
Critical Ratio	W 1.12	P 1.80	W 2.22	W 2.21
Chances in 100 Reliable	W 77	P 89	W 93	W 93
Trials: Mean, Whole Method	3.27	3.09	1.64	2.82
Part Method	6.07	2.61	2.95	4.91
P. E. Mean: Whole Method	0.38	0.45	0.13	0.15
Part Method	0.81	0.14	0.24	0.43
Difference Between Means	W 2.80	P 0.48	W 1.31	W 2.09
P. E. of Difference	0.89	0.47	0.27	0.45
Critical Ratio	W 3.15	P 1.02	W 4.85	W 4.64
Chances in 100 Reliable	W 98	P 75	W 100	W 100

The results of the experiment are summarized in Table III. The expected outcome is given in the first row, and the results from the experiment according to the three measures of time, errors, and trials in order of fineness and probable reliability are listed below. All ratios are those of the difference between means to the probable error of that difference. The results show small but consistent differences in the direction expected with the exception of the two accuracy scores for the first problem.

The disagreement in accuracy scores on the first problem from the general trend may have two causes. (1) The amount of random movement evident in the attack on the first problem was greater than on the subsequent problems, shown in rapid, rather careless activity at first. When the problem was found to be less simple than it at first appeared movements became slower and more methodical. The last three problems did not have this initial period of rapid, careless work. (2) The difference in the direction in which the four segments pointed gave then an element of wholeness which made perception of the rather obscure relationship of the units necessary. Had they all been turned the same way or had any orientation been accepted it is possible the part method would have proved uniformly superior in the first problem.

Intercorrelations for the four problems were worked out on the basis of time scores. It seemed probable that the relationships between problems 1, 2, and 3 would be greater than the relation of any to problem 4 because of the common elements in the design. Table IV shows, however, that the

TABLE IV

INTERCORRELATIONS

Problem	With Problem		
	2	3	4
1	.076 ± .143	.057 ± .144	.067 ± .144
2		− .088 ± .142	− .258 ± .134
3			.091 ± .142

relationships were uniformly low and statistically insignificant. This may be a function of the relatively low reliabilities of the problems, or any relationship existing may have been hidden by combining the results for both methods for the intercorrelations. Method as well as material varied from problem to problem.

In conclusion, results from problems 2, 3, and 4 seem to substantiate the hypothesis that the superiority of the whole method varies roughly with the degree of integration of the unit used. Special factors affecting problem 1 make generalization from that design of doubtful significance.

(Manuscript received January 10, 1936)

A STUDY OF THE RELATIVE AMOUNTS OF FORWARD AND BACKWARD ASSOCIATION OF VERBAL MATERIAL

BY THOMAS G. HERMANS

University of Washington

The reader is doubtless familiar with the controversy regarding the possibilities and relative effectiveness of forward and backward association as carried on by Pavlov, Guthrie, Cason and others. The experimental evidence is varied and extensive. That apparent forward and backward associations do occur is not doubted. This is particularly true in the field of verbal associations. It also is true that with verbal material it is most difficult to determine in which direction the associations have occurred. The problem of the present study is to determine the extent of forward and backward association resulting from the repetition of novel verbal material presented in one sequence only. For example, what would be the result from an association test using as subjects children who had learned the letters of the alphabet in only one sequence? To approximate this hypothetical condition, the present experiment was devised.

Twenty-two nonsense syllables of purported zero association value were selected from the list compiled by Glaze.[1] An attempt was made to select syllables (1) that yield as wide a sampling of the whole alphabet as possible, (2) that are easily pronounced, and (3) that have a minimum rhyming value with others of the list. They were arranged in eleven pairs as follows:

1	2	3	4	5	6	7	8	9	10	11
cef	vec	tov	gah	zum	zif	qub	yem	keb	guk	yop
jic	dax	paf	yud	wez	huc	jid	mib	gex	yof	zil

Each syllable was printed on a card in upper case block type and photographed singly on motion picture film in such an order that each pair was followed and preceded by every other pair once, but within the pair the order was always the same. For example, 'cef' was always followed by 'jic,' but within the series was preceded once with the second one of every other pair; 'jic' was always preceded by 'cef,' but was

[1] Glaze, J. A., The association value of nonsense syllables, *Ped. Sem.*, 1928, 35, 255–267.

followed within the series by the first one of every other pair once. The sequence of pairs were as follows:

1	2	3	4	5	6	7	8	9	10	11
1	3	5	7	9	11	2	4	6	8	10
1	4	7	10	2	5	8	11	3	6	9
1	5	9	2	6	10	3	7	11	4	8
1	6	11	5	10	4	9	3	8	2	7
1	7	2	8	3	9	4	10	5	11	6
1	8	4	11	7	3	10	6	2	9	5
1	9	6	3	11	8	5	2	10	7	4
1	10	8	6	4	2	11	9	7	5	3
1	11	10	9	8	7	6	5	4	3	2

The total series was composed of 220 pictures [2] which could be projected upon the screen as many times as required to effect an appreciable amount of learning on the part of the subject. If the series were presented ten times, the subject would be exposed to the members of a pair in a certain sequence 100 times, but only 10 times to a member of a pair in a sequence relative to any other syllable. If subjects pronounced the syllables as fast as they were flashed upon the screen, it would be expected that associations would be set up between the members of the pair rather than between members of different pairs.

To test the learning, the subject was asked, when the syllables were flashed upon the screen in the test series, to write the first other syllable of the series that occurred to him. The syllables in the test series were arranged in the following order: 2nd syllable of the 1st pair, 1st of the 2nd pair, 2nd of the 3rd pair, 1st of the 4th pair, and so on throughout the twenty-two syllables. Writing the other syllable of the pair was scored as the correct response to the syllable on the screen.

Maximum speed of exposure in the learning series was deemed advisable to prevent opportunity for rehearsal by the subjects between exposure of syllables on the screen. Preliminary experimenting with a few sophisticated subjects indicated that when the projection machine was adjusted to a speed of 75 exposures a minute, it was just possible to read silently each syllable as exposed. With this group also was indicated great difficulty in associating any of the paired syllables, even though the series was presented several times. It then was decided to include in the directions to the naive subjects, the information that (1) the syllables occurred in paired sequence, (2) they were to be given an association test, and (3) the correct association was with the other syllable of the pair. They were instructed to read silently.

Equipped with these directions, a number of subjects in small groups was exposed to the series ten times at the rate of 75 exposures a minute. Short intervals of rest (about two minutes) were allowed after every two repetitions of the series. During the rest period, in an attempt to reduce further the rehearsing factor, the subjects were asked not to think of the syllables and they were distracted by the experimenter who talked to them of such things as the visual fatigue effects peculiar to the conditions of the experiment.

From the test results of those thirty-seven subjects (Group A) who scored one or more correct responses, the following data were derived: of the total 223 associations of paired

[2] Credit is due Professor Edwin R. Guthrie for having devised the photographic and projection technique used in this experiment and I wish to express my gratitude for his assistance in arranging and photographing the syllables.

syllables, 137 or 61.4 percent were forward associations and 86 or 38.5 percent were backward associations. Of this number, however, many points were scored by given individuals on certain pairs in both directions, *i.e.*, they responded to the first syllable of the pair with the second, and to the second of the pair with the first. There were 56 of such points scored in each direction. It was apparent in such cases that the subject had learned the syllables as a pair, and that these data do not serve to differentiate the relative potency of forward and backward association. Tabulating those correct associations which have been formed in one direction only, relative to a pair, we find 81 or 72.9 percent forward, and 30, or 27.0 percent backward associations.

The matter of fatigue proved to be a prominent disturbing factor. Reading syllables for nearly 30 minutes at the rate of 75 a minute with only short rests every six minutes, is no easy task. Some of the subjects reported being unable to attend and read for several consecutive syllables. Others were seen to shut their eyes for rest. It is obvious that during such pauses it is possible for the subject to rehearse the paired syllables. This would tend to vitiate the results obtained. Because of this fatigue, it was decided to change the rate of exposure of the syllables to 60 a minute, thus allowing a more comfortable margin of time for pronouncing completely each syllable. As a result of the change subsequent subjects complained much less of fatigue, although the rate still demanded close attention.

With 89 additional subjects (Group B), the following procedure was used. They were instructed to pronounce silently each syllable as exposed; they were told that the experiment tested some sort of learning, yet they were not to be concerned about that; they were to make of themselves 'Pavlovian dogs' and respond automatically as stimulated and do nothing else. The series was presented eight times with a rest after each even-numbered repetition of the series.

On giving the test only 76 correct associations were scored by the whole group; 39 forward associations and 37 backward. Of this number 34 were cancelled by the subject having

responded to given pairs in both directions. Of the remaining associations in one direction only relative to the pair, 22 or 52.3 percent were forward and 20 or 47.6 percent were backward associations.

The subjects were asked at the end of the test series if they had noticed the fact that the syllables occurred in pairs. About half reported observing at least one pair, but no one observed that all were presented in paired sequence.

This same group of 89 subjects was given further training three days later by an exposure of the series six additional times, making for them in all 14 exposures of the complete series, 140 exposures of each pair. No further directions were given before the training series except the reminder that they were then supposed to know from the previous day's experience that the syllables occurred in pairs, that an association test was to be given at the end of the training, and that they no doubt could guess that they were to learn the pairs.

The association test which followed the training revealed much the same associative tendencies as obtained with Group A. A total of 538 points was scored by the whole group; 303 or 56.3 percent were forward associations, and 235 or 43.6 percent were backward associations. Deducting the points scored by subjects on given pairs in both directions, there were left 172 points, of which number 120 or 69.7 percent were forward and 39 or 30.2 percent were backward associations.

In an attempt to obtain evidence as to the factor of rehearsal of pairs, the subjects in Group B were asked at the end of the test how they learned the pairs—whether they rehearsed, anticipated, or used some other method. Using this information, all of the points scored by those reporting rehearsal of pairs were omitted. The remaining correct responses, after these omissions and the preceding cancellation process, totaled 127, of which number 88 or 69.2 percent were forward, and 39 or 30.7 percent were backward associations.

Another way of handling the test results appeared to promise further evidence bearing upon the problem. Each syllable of a pair was preceded or followed in the training

series by syllables of other pairs. If a subject responded associatively to the first of a pair by writing the second of some other pair, it could be regarded as backward association. Responding to the second of a pair with the first of some other pair could be regarded as forward association. On tabulating these responses it was found that 56 percent were backward associations and the remainder forward associations. This fact, although at first startling, has little significance in our present problem. Upon consideration of the individual syllables associated it became obvious that the associations with other syllables of the series was attributable to similarities of sound, appearance or other undetermined associative values.

The question immediately arose, however, as to how much of the forward or backward associations of the syllables arbitrarily paired was due to undetermined associative values which may be a predisposing factor toward associations in one direction rather than the other, in considering the gross results of the tests. The use of eleven pairs did not seem to be a sufficient number to guarantee the cancelling out of such a factor if it were involved.

To check this factor, a group of 60 subjects was exposed to the training series presented in reverse order. Approximately the same relative amounts of forward and backward associations were found as were found with the former groups. From this check it is obvious that the relative amounts of forward and backward associations of the paired syllables are not attributable to the chance and undetermined association values of the particular syllables paired.

Discussion of results and conclusion.—The main conclusion that may be drawn from the foregoing experiment is that apparent backward association does occur in the learning of verbal material. The word 'apparent' is used here because the significance of the results obtained is not at all clear. It is impossible to control and determine the actual sequence of verbal responses that are elicited simply through the relation of the pairs presented upon the screen. In other words,

responding to the stimulus of the second of the pair with the first of the pair is backward association only in consideration of the objective pair as presented during training.

There are two reasons for questioning the significance of the results. First, in learning these paired nonsense syllables the subject can rehearse the pair, and in so doing, link them together associatively in a manner that may be either forward or backward or both. The fact that there was no change in the relative amounts of apparent forward and backward associations when the scores of those reporting rehearsal were eliminated, does not preclude the possibility, some might say, of unconscious rehearsal by those yielding the remaining data. Yet if this were a factor, one would expect some reduction in the percent of backward association upon such elimination. One might take arbitrarily as a criterion of rehearsal, whether conscious or unconscious, the fact that a subject scored points in both directions relative to a pair. This was done; all of the points scored by any subject who had made such associations on both members of a pair were cancelled. The remaining data from both tests of Group B and the one test of Group A, yielded 89 scored points, 39.3 percent of which were backward associations. The inclusion of data from both tests of Group B is legitimate since practically every one whose first test performance yielded data, was excluded on the present basis of selection by his performance on the second test. However, excluding the data from the first test of Group B there is still found that of the 51 points remaining, 36.6 percent were backward associations. It is apparent that rehearsal by this criterion also does not account for the showing of backward association obtained.

The second reason for questioning the significance of the results obtained is that the apparent backward association may be attributable to association of the simple simultaneous variety. Upon responding to the first of the pair in the training series, the response may be maintained or repeated until the visual stimulus of the second is given. In such an event the visual stimulus of the second syllable may come to set off this response to the first. That this did occur fre-

quently enough to account for the showing of the apparent backward associations obtained is very doubtful. Most of the subjects reported no difficulty in following the instruction to pronounce silently each syllable as flashed on the screen. This meant that their response to the second of the pair was appropriate to it, and not a response to the first of the pair.

The conclusion to be drawn from the foregoing experiment is that backward associations do occur in learning verbal material.

(Manuscript received November 16, 1935)

DISTRIBUTION OF EFFORT IN A PERCEPTUAL-MOTOR TASK. SERIES II

BY DOROTHY RANSOM AND E. B. SKAGGS[1]

Wayne University

The first study, Series I, of the effect of distribution of effort was reported in 1935.[2] Since the present and earlier study were done under relatively the same conditions of procedure, we shall refer the reader to a description of the apparatus and procedure as reported in detail in the earlier study. However, enough will be said here to give the reader, who is not acquainted with the earlier report, a general idea of the 'set-up' and procedure used.

The apparatus consisted of two Goddard formboards fastened to the top of a table, one near, the other farther away from, the subject. The task consisted in replacing the blocks from the more distant board to their proper recesses in the nearer board. Time was taken in seconds and fractions thereof. Hence, all figures which appear in the following paper express seconds or fractions thereof. Efficiency is always the greater the smaller the figure.

Certain differences between this and the earlier experiment must be noted. *First*, Miss Ransom, a graduate student, was experimenter in this investigation whereas the senior author, a man, and the subjects' instructor, was experimenter in the first study. *Second*, every subject worked the board, in all, 122 times (trials) in the present investigation, whereas 112 trials were completed in the earlier study. *Third*, the subjects in the present study were sophomore men while in the earlier study they were sophomore women.

[1] Miss Ransom carried through this experiment as part of her work for the M.A. degree. She was specifically trained to give the tests and record times by the senior writer just as he did earlier. All computations and conclusions are by the senior writer.

[2] Skaggs, E. B. Distribution of effort in a perceptual-motor task. *J. Exper. Psychol.*, 1935, **18**, 797–801.

Fourth, the method of testing final efficiency differed somewhat from that used in the earlier study. Regardless of method of distribution of effort used, every subject finished the regular practice work on a Friday. On the following Monday every subject returned to the laboratory to take the final 'efficiency test' of ten trials. In the earlier study this final and extra session of ten trials was not held. Efficiency on the last ten trials was used, but the last ten trials, in the case of Group A, were the final 10 of the last 20 trials, and, in the case of Group C, a combination of the 5 trials of the last two meetings. Obviously, the procedure followed in the present study is more desirable in many ways.

As in the earlier study, three groups of subjects were used. Group A, 20 subjects, worked the board 20 times every day for 5 days. Group B, 19 subjects, worked the board 10 times every day for 10 days, omitting Saturday and Sunday. Group C, 18 subjects, worked the board 5 times every day for 20 days, omitting Saturdays and Sundays.

Every subject, regardless of group, first reported on a Friday for instruction in doing the task and watching the experimenter work it. At that time, after two trials in which no time was taken, every subject worked the board ten trials. The average time of these initial 10 trials was taken as the index of 'initial formboard ability.' Regular work began on the following Monday. As stated above, every subject, regardless of group, returned after finishing the regular practice period of 100 trials, for a final test of 10 trials. Likewise, four weeks later every subject returned to take a retention test of 10 trials. All knew beforehand that this four-week retention test would be made.

As in the earlier study, our subjects were college sophomores of about the same chronological age, ranging between 18 and 22 years, with the median about 20 years. They were of approximately the same intelligence level as measured by the Otis Self-Administering Intelligence Tests, Higher Form, 20-minute work-time. Their scores (group averages) were as follows: Group A, 47.8 (range 35–68); Group B, 51.2 (range 35–71); and Group C, 52.4 (range 32–74). Group A averaged 3 points below B and 4 points below C.

Initial formboard ability was about the same for the three groups as is shown by the figures of Table I. Group A seems to be slightly slower, on the average, however, than the other two groups. If one judges the relative learnability of the groups by comparing the amount of gain of the first so-called 'regular' practice trials over the first 10 practice trials one notes (Table I) that A gains 0.83 or 15 percent, B gains 0.53 or 14.4 percent, and C gains 0.57 or 11 percent. According to these comparisons Groups A and B are about equal in 'learnability' in this task while Group C is somewhat inferior, although the group has the highest intelligence rating.

TABLE I

PRESENTING A COMPARISON OF INITIAL FORMBOARD ABILITY OF GROUPS A, B, AND C, BASED UPON THREE CRITERIA

Criterion Used		Group A	Group B	Group C
First 10 practice trials *	Mean	5.56	5.11	5.15
	M.V.	0.46	0.43	0.48
	S.D.	0.59	0.59	0.65
First 10 regular trials	Mean	4.73	4.58	4.58
	M.V.	0.49	0.36	0.34
	S.D.	0.59	0.44	0.43
First 20 regular trials	Mean	4.53	4.48	4.43
	M.V.	0.41	0.27	0.45
	S.D.	0.51	0.33	0.52

* Criterion stressed in conclusions.

TABLE II

PRESENTING A COMPARISON OF GROUP AVERAGES AND MEASURES OF VARIABILITY FOR GROUPS A, B, AND C

Criterion Used		Group A	Group B	Group C
Last 20 trials	Mean	3.52	3.63	3.56
	M.V.	0.23	0.23	0.26
	S.D.	0.29	0.29	0.32
Last 10 trials *	Mean	3.66	3.59	3.50
	M.V.	0.29	0.23	0.30
	S.D.	0.34	0.28	0.36
Last 5 trials	Mean	3.62	3.55	3.46
	M.V.	0.29	0.28	0.35
	S.D.	0.34	0.33	0.42

* Criterion stressed in conclusions.

Results

If one overlooks the small differences shown in Table I and regards our three groups as having equal formboard ability, then the data of Table II, which presents figures of final ability, may be taken as absolute values. As such they show that, in the case of the last 10 and last 5 trials, greater efficiency comes with greater distribution of effort. We have stressed the use of the last 10 trials as the best criterion for judging final ability. The reasons for this are fairly obvious.

TABLE III

Showing Group Averages and Measures of Variability on the Four-Week Retention Test

	Group A	Group B	Group C
Mean	3.71	3.46	3.63
M.V.	0.28	0.22	0.24
S.D.	0.35	0.26	0.28

It may be objected at once, however, that, since the initial scores are not absolutely alike, one must compare efficiency of our groups in terms of the *relative amount of gain* of final over initial trials. Table IV presents these relative gains. According to the figures of Table IV Group A, with least distribution of effort, stands intermediate between Groups B and C. One notes that A has gained more than C. Referring back to Table I and the discussion which just precedes that Table, it will be noted that Group C seems to be about 4 percent poorer learners than A or B. This difference in learnability might take care of the final differences between A and C, a difference of 2 percent. However, this does not care for the difference between Groups B and C.

TABLE IV

Comparing Absolute and Relative Amount of Gain for Each Group on Final 10 Trials Over Initial 10 Trials

	Group A	Group B	Group C
Mean, first 10 trials	5.56	5.11	5.15
Mean, final 10 trials	3.66	3.59	3.50
Gain	1.90 or 34%	1.52 or 43%	1.65 or 32%

All in all, one may argue for a tendency for the more distributed forms of effort to yield greater efficiency, but certainly the tendency is not as clear-cut as one would wish. One scarcely dares to make a definite statement one way or the other.

Tables III and V present the results of the efficiency test

TABLE V

COMPARING ABSOLUTE AND RELATIVE AMOUNT OF GAIN OR LOSS FOR EACH GROUP ON 4-WEEK RETENTION TEST OVER FINAL 10 TRIALS

	Group A	Group B	Group C
Mean, 4-week retention test	3.71	3.46	3.63
Mean, final 10 trials	3.66	3.59	3.50
Gain (+) or loss (−)	− 0.05 or 1.4%	+ 0.13 or 3.6%	− 0.13 or 3.7%

made exactly four weeks after the final efficiency test at the end of the practice period. Under the assumption that the groups were of equal initial ability, despite the small differences shown in Table I, the data indicate that Group B is most efficient and Group A least. But again it may be said that only relative figures are of any value. Hence, Table V is presented. According to this Table, Group B *gains* approximately the same amount that Group C *loses*, over the final efficiency score. It is interesting to note that Group B actually gains in efficiency after being away from the task for four weeks! Group C, with greatest amount of distribution, loses more than Group A, with least amount of distribution of effort in training.

Again the data are not at all clear in their indication. We do not feel justified in saying that the greater distribution of effort resulted in least forgetting and *vice versa*.

COMPARISON OF RESULTS OF PRESENT AND EARLIER STUDY

In the earlier study with women subjects, Group A improved 37 percent on the final 10 trials over the initial 10 trials; Group B gained by 32.5 percent; and Group C improved 34 percent. Even granting that Group A was constituted of better learners than Group C, since A gained 13.5 percent and C only 11.6 percent in the first 10 'regular trials' over the first

10 'practice trials,' there still remained a residue of greater gain to be accredited to Group A. Groups A and B were equal in learnability and yet Group B improved less than A. Making comparisons in terms of relative gains or losses in the 4-week retention tests over the final ten practice trials, little can be said concerning the relation of distribution of effort and ability to retain in the task.

On the whole, the trend of the earlier experiment was toward a greater efficiency of the more massed method of learning. Certainly, the most astute critic of the data presented could find nothing to support the view that distributed effort, as used in that experiment, was superior.

The conditions, as stated at the beginning of this paper, were not identical with those of the earlier experiment. Hence, the data are only roughly comparable. In the present experiment we have again failed to find any clean-cut evidence of the superiority of the more distributed form of practice.

The writer suggests the opinion that, in both of the investigations referred to in this paper, the distributed forms of procedure have suffered from the fact that the students who acted as subjects did not keep up a uniformly high degree of interest and effort. In their reports they all said that they did keep up a high degree of interest and effort. However, the writers cannot be so naïve as to take these reports at face value. After all, it is asking much, indeed, of an ordinary undergraduate student, whose laboratory work is compulsory, to come regularly at a definite appointed time to the laboratory for two weeks or four weeks to work a form-board. True enough, we tried to encourage these students and interest them in the work, but did we succeed equally with all groups? We doubt it. All of this work needs to be repeated with the presence of strong incentives (as worthwhile sums of money-rewards for relative individual improvement) which will really insure a sustained interest and effort. If, on the other hand, we are incorrect in our suspicion of unequal effort and striving on the part of our groups, then the failure to find a definite superiority of the more distributed form of practice must find some other explanation.

Conclusion

In the present investigation, where a perceptual-motor task has been employed with three different forms of distribution of practice, we have failed to find any clear-cut evidence of the superiority of the more distributed form of practice. This was likewise true of the earlier investigation.

It is suggested that in both investigations it is possible that, with the more distributed forms of practice, the subjects did not keep as sustained interest, effort and striving as they did in the more massed form of practice, in which the experiment 'was ended' quicker.

An interesting by-product of both investigations has been the fact that, with the degree of over-learning involved, there has been almost no loss over a period of four weeks and, in some cases, an actual increase in efficiency.

(Manuscript received January 15, 1936)

THE NEW LABORATORY OF PSYCHOLOGY AT THE UNIVERSITY OF ROCHESTER

BY LEONARD CARMICHAEL

University of Rochester

At the time of the erection of the new River Campus of the University of Rochester in 1930 which was largely made possible by the munificence of the late Mr. George Eastman, a psychological laboratory was provided in Morey Hall. During the present year, however, this laboratory has been much enlarged in order to provide facilities for instruction and research on both the undergraduate and graduate levels in human psychology, physiological psychology, comparative psychology, and the psychology and physiology of the sense organs.

FIG. 1. Laboratory of Psychology, University of Rochester.

The new laboratory comprises two entire floors of a building approximately 209 by 45 feet. The building is of heavy, sound-resisting, fireproof construction throughout. The ceilings are approximately 13 feet high.

All the rooms of the laboratory are supplied with direct and indirect electric current. A switchboard provides for the electrical inter-connection of a number of the rooms on special circuits. Gas, compressed-air, and vacuum are available in certain of the rooms.

A simplified plan of the laboratory is given as Fig. 1. In the following paragraphs, a brief description is presented of the individual rooms of the laboratory and of their special features.

Floor B

Room 100—Student Laboratory for General Experimental Psychology.—This room is fitted with a series of instrument cases with glass doors. It contains seven large tables, each so arranged that electrical outlets are directly available. There is also a large demonstration desk in this room fitted with a variety of electrical outlets, gas, and a sink. The dimensions of this room are approximately 40 by 35 feet.

Room 101—Vivarium and Preparation Room.—This well-lighted room is separated by a solid partition into two sections, each of which is separated from an inner hallway by a heavy wire mesh partition with locked doors. One of the inner rooms is fitted with tiers of cages including specially constructed all-metal cages for cats and monkeys; the other room contains all-metal racks and cages for guinea-pigs and rats. These cages and racks are welded, not soldered, and the whole arrangement is such that cleaning is facilitated. A large preparation table, bins for the storage of food, sawdust and the like, two large sinks, and several metal cabinets for the storage of supplies are also provided in this room. It is ventilated by a special internal system such that the outside windows need never be opened. Special steam lines and suitable thermostats make it possible to maintain the temperature of the vivarium at any desired level, with a fluctuation of only a few degrees. The dimensions of this room are approximately 26 by 22 feet.

Rooms 102 A and B—Offices or Research Rooms.—At present these rooms will be used as offices for graduate assistants. Each of these rooms is approximately 10 by 10 feet.

Rooms 103, 104 and 106—Offices.—These rooms provide offices for members of the departmental staff. Each room is approximately 10 by 20 feet.

Room 105—Staff Research in Physiological Psychology.—This room, painted in white enamel throughout, contains an operating table, instrument cases, a pressure autoclave and a sink with foot-controlled faucets. The bath apparatus used in research in fetal behavior is located here. All facilities of this room have been developed to make possible completely aseptic operations on animals. This room also contains a research histological laboratory including a new Bausch and Lomb Minot-type sliding microtome for brain sections. The operating tables of this room are so placed that they receive a large flood of direct daylight. There is also provision for effective artificial lighting over the operating and research tables. A special arrangement for the motion-picture recording of fetal behavior and of operative procedures has also been provided in this room. The approximate dimensions of the room are 25 by 21 feet.

Room 108—Lecture Room—Classes in Psychology.—This room is provided with a large demonstration lecture table with electrical and other outlets useful in general demonstrations. The dimensions of this room are approximately 34 by 30 feet.

Rooms 109 and 110—Research Rooms.—Fully dark, sound-resisting, each approximately 11 by 14 feet.

Room 111—Photographic Dark Room.—This room is entered through a 'light trap.' The room is fitted with a double soapstone sink for photographic work. It also contains storage cabinets and special apparatus for the development of long strips of recording paper. The dimensions of this room are approximately 10 by 16 feet.

Room 112—Research Room.—Fitted with inner dark chamber especially devised for after-image work in vision. This room is approximately 9 by 14 feet.

Room 113—Apparatus and Stock Room.—Storage batteries and large pieces of apparatus not contained in cases in Room 100 are stored here. The dimensions of this room are approximately 12 by 11 feet.

Rooms 114 A and B—Laboratory Shop and Preparation Room.—These rooms are fitted with benches and hand tools for metal and wood work. The headquarters of the full-time laboratory mechanician are located here, although he uses other university shops for work requiring power machines. The dimensions of these two connecting rooms are approximately 25 by 11 feet.

Floor A

Room 200—Drafting Room.—This room contains suitable lights, a large drafting table and instruments for making charts and diagrams. The dimensions of the room are approximately 8 by 14 feet.

Rooms 200 A, B, C and D—Offices for Members of the Departmental Staff, Assistants and Graduate Students.—These rooms vary in size but each is approximately 8 by 10 feet.

Room 201—Seminar Room and Departmental Library.—The walls of this room are fitted with book shelves except for one side on which a large blackboard is provided. The room is so planned that entrance and exit to it may under certain circumstances be only through the adjoining room 202 A. This arrangement provides a means of supervising the charging of books borrowed from the library. The dimensions of the room are approximately 36 by 22 feet.

Room 202 A—Office of the Departmental Secretary.—As noted above, entrance to the library, as well as to the office of the director of the laboratory, is through this room. The portion of this room containing the secretary's desk, files, and so forth is separated from the rest of

the room by a low partition. In the outer portion of the room are chairs for those waiting to enter the offices. A pigeon-hole box for the distribution of laboratory mail and a laboratory notice-board are located here. The dimensions of this room are approximately 10 by 20 feet.

Room 202—Office of the Director of the Laboratory.—The dimensions of this room are approximately 18 by 20 feet.

Room 203 A—Staff Office.—The dimensions of this room are approximately 10 by 20 feet.

Room 203—Student Laboratory of Comparative Psychology.—This room is fitted with suitable tables and a demonstration bench. It is used for permanently organized pieces of apparatus employed in instruction in comparative psychology. The dimensions of this room are approximately 18 by 20 feet.

Room 204—Statistical Laboratory.—This room is administered by the Department of Economics, but it is also used by students in the Department of Psychology. This room is fitted with a variety of electrical calculating machines and other statistical devices. Its dimensions are approximately 36 by 32 feet.

Room 205—Sound-proof Room.—Work of sound-proofing this room is not yet complete but, following the plans kindly provided by Dr. J. F. Shepard of the University of Michigan, it is hoped to make this a fully satisfactory sound-proof room. This room will also be electrically shielded. Its outside dimensions are approximately 20 by 12 feet.

Room 206 A—Electrically Shielded Room.—This room is entirely shielded inside with soldered bronze screen of very fine mesh. It is at present equipped with a Westinghouse oscillograph and especially constructed amplifiers. This apparatus makes possible the photographic recording of many types of bio-electrical phenomena. The apparatus in its present form is especially suitable for recording the Berger rhythm of the human cortex and for studies of the electrical phenomena accompany-

ing behavior in young and fetal organisms. Connection between this room and the sound-proof room is provided for. The dimensions of this room are approximately 12 by 16 feet.

Rooms 206 B, C and D—Research Rooms.—Dark, sound-resisting, each approximately 12 by 12 feet.

Rooms 207 and 208—Laboratory Rooms.—May be used either for research or for instruction in advanced laboratory courses. The dimensions of each room are approximately 18 by 12 feet.

The whole laboratory is connected by tunnels with the general university library and with the laboratories of biology, chemistry, physics and physiological optics, and with the general university shops including the optics shop. Direct connection is also available with the library and laboratories of the medical school.

Coincident with the opening of the new laboratory, three new graduate fellowships, six new graduate assistantships and two scholarships in psychology have been announced by the University of Rochester. All these appointments are now filled for the current year.

Graduate work in psychology at Rochester will also be carried on in close association with other research departments of the Faculty of Arts and Science, especially including the departments of the biological sciences, physiological optics and with certain departments of the medical school. Work in the psychology of music carried on in the Eastman School of Music of the University of Rochester is directly connected with the department.

In Carnegie Hall on the campus of the Women's College of the University of Rochester, an excellent laboratory for undergraduate work in psychology is also maintained under the direction of the department.

(Manuscript received September 21, 1936)

PSYCHOLOGICAL REVIEW PUBLICATIONS

Original contributions and discussions intended for the Psychological Review should be addressed to

Professor Herbert S. Langfeld, Editor PSYCHOLOGICAL REVIEW,
Princeton University, Princeton, N. J.

Original contributions and discussions intended for the Journal of Experimental Psychology should be addressed to

Professor Samuel W. Fernberger, Editor JOURNAL OF EXPERIMENTAL PSYCHOLOGY,
University of Pennsylvania, Philadelphia, Pa.

Contributions intended for the Psychological Monographs should be addressed to

Professor John F. Dashiell, Editor PSYCHOLOGICAL MONOGRAPHS,
University of North Carolina, Chapel Hill, N. C.

Reviews of books and articles intended for the Psychological Bulletin, announcements and notes of current interest, and *books offered for review* should be sent to

Professor John A. McGeoch, Editor PSYCHOLOGICAL BULLETIN,
Wesleyan University, Middletown, Conn.

All business communications should be addressed to

Psychological Review Company, Princeton, New Jersey

The

Journal of Experimental Psychology

is indexed in the

International Index to Periodicals

to be found in most public and college libraries

AMERICAN PSYCHOLOGICAL PERIODICALS

American Journal of Psychology—Ithaca, N. Y.; Cornell University.
Subscription $6.50. 624 pages annually. Edited by K. M. Dallenbach, M. F. Washburn, Madison Bentley, and E. G. Boring.
Quarterly. General and experimental psychology. Founded 1887.

Journal of Genetic Psychology—Worcester, Mass.; Clark University Press.
Subscription $14.00 per yr.; $7.00 per vol. 1,000 pages ann. (2 vols.). Edited by Carl Murchison.
Quarterly. Child behavior, animal behavior, comparative psychology. Founded 1891.

Psychological Review—Princeton, N. J.; Psychological Review Company.
Subscription $5.50. 540 pages annually. Edited by Herbert S. Langfeld.
Bi-monthly. General psychology. Founded 1894.

Psychological Monographs—Princeton, N. J.; Psychological Review Company.
Subscription $6.00 per vol. 500 pages. Edited by John F. Dashiell.
Without fixed dates, each issue one or more researches. Founded 1895.

Psychological Index (discontinued after 1936)—Princeton, N. J.; Psychological Review Co.
Subscription $4.00. 400–500 pages. Edited by Walter S. Hunter and R. R. Willoughby.
An annual bibliography of psychological literature. Founded 1895.

Psychological Bulletin—Princeton, N. J.; Psychological Review Company.
Subscription $6.00. 720 pages annually. Edited by John A. McGeoch.
Monthly (10 numbers). Psychological literature. Founded 1904.

Archives of Psychology—New York, N. Y.; Columbia University.
Subscription $6.00. 500 pages per volume. Edited by R. S. Woodworth.
Without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal and Social Psychology—Eno Hall, Princeton, N. J.; American Psychological Association.
Subscription $5.00. 448 pages annually. Edited by Henry T. Moore.
Quarterly. Abnormal and social. Founded 1906.

Psychological Clinic—Philadelphia, Pa.; Psychological Clinic Press.
Subscription $3.00. 288 pages. Edited by Lightner Witmer.
Without fixed dates (Quarterly). Orthogenics, psychology, hygiene. Founded 1907.

Journal of Educational Psychology—Baltimore: Warwick & York.
Subscription $6.00. 720 pages. Monthly except June to August.
Edited by J. W. Dunlap, P. M. Symonds and H. E. Jones. Founded 1910.

Psychoanalytic Review—Washington, D. C.; 3617 10th St., N. W.
Subscription $6.00. 500 pages annually. Edited by W. A. White and S. E. Jelliffe.
Quarterly. Psychoanalysis. Founded 1913.

Journal of Experimental Psychology—Princeton, N. J.; Psychological Review Company.
Subscription $7.00. 900 pages annually. Edited by Samuel W. Fernberger.
Bi-monthly. Experimental psychology. Founded 1916.

Journal of Applied Psychology—Indianapolis; C. E. Pauley & Co.
Subscription $6.00. 600 pages annually. Edited by James P. Porter, Ohio University, Athens, Ohio. Bi-monthly. Founded 1917.

Journal of Comparative Psychology—Baltimore, Md.; Williams & Wilkins Company.
Subscription $5.00 per volume of 450 pages. Ed. by Knight Dunlap and Robert M. Yerkes. Two volumes a year. Founded 1921.

Comparative Psychology Monographs—Baltimore, Md.; The Johns Hopkins Press.
Subscription $5.00. 400 pages per volume. Roy M. Dorcus, Editor.
Published without fixed dates, each number a single research. Founded 1922.

Genetic Psychology Monographs—Worcester, Mass.; Clark University Press.
Subscription $7.00 per vol. One volume per year. Edited by Carl Murchison.
Bi-monthly. Each number one complete research. Child behavior, animal behavior, and comparative psychology. Founded 1925.

Psychological Abstracts—Eno Hall, Princeton, N. J.; American Psychological Association.
Subscription $6.00. 700 pages ann. Edited by Walter S. Hunter and R. R. Willoughby.
Monthly. Abstracts of psychological literature. Founded 1927.

Journal of General Psychology—Worcester, Mass.; Clark University Press.
Subscription $14.00 per yr.; $7.00 per vol. 1,000 pages ann. (2 vols.). Edited by Carl Murchison.
Quarterly. Experimental, theoretical, clinical, historical psychology. Founded 1927.

Journal of Social Psychology—Worcester, Mass.; Clark University Press.
Subscription $7.00. 500 pages annually. Ed. by John Dewey and Carl Murchison.
Quarterly. Political, racial, and differential psychology. Founded 1929.